LSHTM Research Online

Goodman, A; (2009) Why do British Indian children have an apparent mental health advantage? PhD thesis, London School of Hygiene & Tropical Medicine. DOI: https://doi.org/10.17037/PUBS.00942773

Downloaded from: https://researchonline.lshtm.ac.uk/id/eprint/942773/

DOI: https://doi.org/10.17037/PUBS.00942773

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license. To note, 3rd party material is not necessarily covered under this license: http://creativecommons.org/licenses/by-nc-nd/3.0/

https://researchonline.lshtm.ac.uk

Why do British Indian children have an apparent mental health advantage?

Anna Goodman

London School of Hygiene and Tropical Medicine
University of London

Thesis submitted for the degree of PhD, September 2009

Part 1: Main body of thesis

Declaration of authorship

I, Anna Goodman, declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

Signature: ______________________________

Date: 8th December 2009

Word count (exclusive of appendices and bibliography): 99 561 words

Funding and ethical approval

This research degree was funded by an ESRC/MRC studentship. It received ethical approval from the LSHTM ethics committee on 20/12/2007 (application number 5027).

Acknowledgements

I am very grateful to Rebecca Gatward and other colleagues at the Office for National Statistics for their cooperation in accessing the data for this study, and to colleagues at the Office for Standards in Education (OFSTED) for their help in creating a 'Ford score' for schools in B-CAMHS04. I would like to thank Pablo Mateos and Juliana Cipla for allowing me to use a pilot version of the name-matching program Onomap, and to thank Robert Goodman, Sophie Plagerson, Marta Vazquez Ramallo and Kristina Stanfield for their assistance in independently cross-checking my study inclusion and data extraction in my systematic review. I am also very grateful to Simon Cousens, Judy Green, Donna Lamping and Paul Wilkinson for sitting on my advisory committee, and to Mike Kenward and George Ploubidis for providing advice.

Finally, many thanks to my supervisors Dave Leon and Vikram Patel for their encouragement and support throughout this PhD.

Abstract

The British Child and Adolescent Mental Health Surveys (B-CAMHS) of 1999 and 2004 found a substantially lower prevalence of any child mental disorder in Indians compared to the general population (3.4% vs. 9.4%, $p<0.001$). This PhD sought to understand this apparent Indian mental health advantage through secondary analyses comparing the 16 449 White and 419 Indian children aged 5-16 in B-CAMHS.

There was strong evidence ($p<0.002$) of an Indian advantage for externalising problems/disorders and little or no difference for internalising problems. This was consistently observed for clinical diagnosis and for the Strengths and Difficulties Questionnaire (SDQ) administered separately to parents, teachers and children. Detailed psychometric analyses provided no evidence that measurement bias could account for this observed Indian advantage. There was likewise no evidence that the advantage could be explained by participation bias.

In multivariable analyses the unexplained difference between Indians and Whites for externalising problems decreased somewhat after adjusting for the fact that Indian children were more likely to live in two-parent families (92.2% vs. 65.4%) and less likely to have academic difficulties (e.g. 2.9% vs. 8.6% for parent-reported learning difficulties). In models adjusting for a larger number of child, family, school and area variables the difference reduced only by about a quarter (e.g. from 1.08 to 0.75 SDQ points on the parent SDQ) and remained highly significant ($p<0.001$). In both unadjusted and adjusted models, the unexplained Indian advantage for externalising problems was consistently larger in families of low SEP. There was little or no evidence of an ethnic difference for internalising problems/disorders in unadjusted or adjusted models.

In conclusion, the Indian mental health advantage is genuine and is specific to externalising problems/disorders. Family type and academic abilities mediate part of this advantage, but most of the advantage is not explained by major child mental health risk factors.

Table of contents

Tables in thesis

Figures in thesis

Glossary of abbreviations used in this thesis

ADHD: Attention deficit hyperactivity disorder.

ANOVA: Analysis of variance

ASEBA: Achenbach System of Empirically Based Assessment

AUC: Area under the curve (used in ROC analyses)

B-CAMHS: The British Child and Adolescent Mental Health Surveys, for 1999 (B-CAMHS99) and 2004 (B-CAMHS04).

BAS-II: British Ability Scales, second edition ability.

BPVS-II: British Picture Vocabulary Scale, second edition

BCS: British Cohort Study

CBC: The Child Benefit Centre

CI: Confidence interval

CFI: Comparative Fit Index

CFA: Confirmatory factor analyses

CMD: Common mental disorders

DAWBA: Development and Well-Being Assessment.

DSH: Deliberate self harm

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition

EAT-26: Eating Attitudes Test-26 item

EFA: Exploratory factor analyses

FAD: Family Assessment Device

GAD: Generalised anxiety disorder

GF scale**:** General Functioning scale (subscale of the Family Assessment Device)

GHQ-12: General Health Questionnaire, 12-item version

GLM: Generalised linear model

HALS: Health and Lifestyle Survey

HSE: Health Survey for England

ICC: Intra-class correlations

ICD-10: International Statistical Classification of Diseases, version 10

IID: Independently identically distributed

IMD: Indices of Multiple Deprivation

LSOA: Lower Super Output Area

MAR: Missing at random

MCAR: Missing completely at random

MICE: Multiple imputation by chained equations

MLE: Maximum likelihood estimation.

MTMM: Multitrait-multimethod analyses

NCDS: National Child Development Study

NHS: National Health Service

NMAR: Not missing at random

NPV: Negative predictive value

NOS: Not Otherwise Specified

NS-SEC: the National Statistics Socio-Economic Class classification system

OCD: Obsessive-compulsive disorder

ODD: Oppositional defiant disorder

OFSTED: The Office for Standards in Education.

ONS: The Office of National Statistics.

PCA: Principal component Analysis

PML: pseudo maximum likelihood

PPV: Positive predictive value

PTSD: Post traumatic stress disorder

RMSEA: Root Mean Square Error of Approximation

ROC analyses: Receiver operating characteristic analyses

SAS: Social Aptitudes Scale

SDQ: Strengths and Difficulties Questionnaire

SEP: Socio-economic position

SOC: Registrar General's Standard Occupational Classification system

TDS: Total difficulties score [on the SDQ]

TLI: Tucker Lewis Index

WHO: World Health Organisation

WLSMV: Weighted Least Squares, mean and variance adjusted estimator

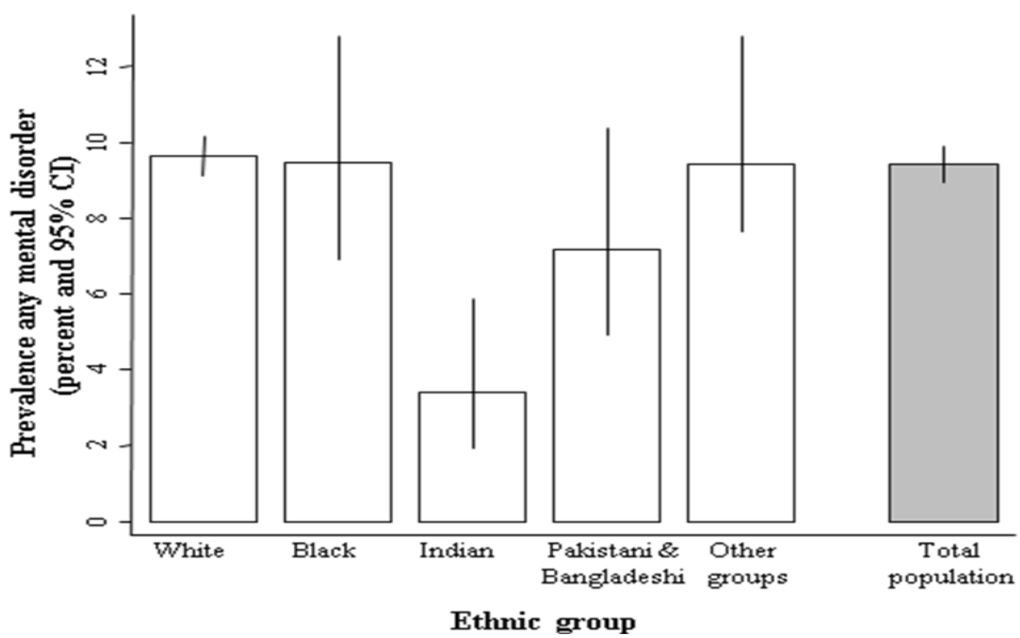
Chapter 1 Introduction

1.1 Rationale

1.1.1 A ‘mystery story’

This PhD falls into the research genre of a ‘mystery story’ [1] – that is, it starts with an observed research puzzle and tries to explain it. The research puzzle in question is the much lower prevalence of mental disorders diagnosed in British Indian children and adolescents in the British Child and Adolescent Mental Health Surveys (B-CAMHS) of 1999 [2] and 2004 [3]. These were two cross-sectional studies of child mental health in England, Scotland and Wales which took nationally-representative samples from the general population. The two surveys achieved a total combined sample of 18 415 children aged 5-16. Figure 1.1 displays the prevalence of any mental disorder by ethnic group, using the ethnic groupings employed by the B-CAMHS reports. In Indians the estimated prevalence of any mental disorder is 3.4% (95%CI 1.9%, 5.9%), substantially lower than all other ethnic groups presented and also lower than the general population prevalence of 9.4% (95%CI 9.0%, 9.9%).

Figure 1.1: Prevalence of any mental disorder by ethnicity in B-CAMHS99 and B-CAMHS04 combined (five-way classification of ethnicity)



1.1.2 Why this mystery might matter

The B-CAMHS surveys therefore suggest a marked mental health advantage for Indian children relative to other ethnic groups (for brevity, I abbreviate 'British Indian' to 'Indian' throughout this PhD, and use 'children' to include adolescents). This is not a trivial observation. Child mental disorders are common around the world, with most population-based surveys producing point prevalence estimates of 10-20% [4]. These rates are high in absolute terms, and also represent a disproportionately large percentage of the burden of disease in children as compared with adults [5]. In Britain, the mental health of children has deteriorated in the past 50 years [6-7], even while their physical health has generally improved. In this context, the centrality of mental health to child welfare has increasingly been recognised by academics, health service providers and policy-makers [8]. Knowing why this one particular group of British children appears to be doing so well could therefore hold important clues to improving the well-being of children from all ethnic groups.[1]

Yet relatively little is known about the epidemiology of social and cultural variables like ethnicity in relation to child mental health. No detailed work on this topic has been done using B-CAMHS data, and several reviews of the literature have commented on the paucity of evidence on minority ethnic children [10-13]. This has been noted to apply particularly to South Asians (i.e. Indian, Pakistani, Bangladeshi and Sri Lankan children). These reviews also revealed that while several studies indicate lower rates of service use in South Asian children, the cause of this difference is unclear [10, 13].

Previous evidence therefore provides little basis for evaluating whether the apparent Indian mental health advantage in Figure 1.1 reflects a real health difference – that is, whether mental health problems are truly less frequent in Indian children. This cannot simply be assumed: an important alternative interpretation is that the difference results from under-recognition or under-reporting of problems in Indian children. From a public health

[1] See 9. Patel, V. and A. Goodman, *Researching protective and promotive factors in mental health.* Int J Epidemiol, 2007. **36**(4): p. 703-7. (Appendix 3) for an editorial in which Vikram Patel and I argue this point more generally with regard to studying advantaged groups.

perspective this distinction is clearly important. Real health differences may hold important clues for aetiology and prevention, while the implications of reporting biases mainly relate to issues of detection and referral.

These considerations inspired me to use this PhD to investigate why Indians show this apparent mental health advantage. Clarifying between competing explanations will make an important contribution to our understanding of the mental health of Indian children, and may generate insights about the aetiology of child mental health more generally. Moreover, through a detailed analysis of this particular cross-cultural difference, I hope to contribute to the development of methodological tools with broader application in cross-cultural psychiatric epidemiology.

A further motivation for my choice of research topic was an awareness that epidemiological research into the health of minority ethnic groups often ignores diversity between minority groups and adopts a problem-centred attitude towards minority ethnic health ([14-15]; see also Chapter 3). Likewise most psychiatric epidemiology focuses upon risk factors and high-risk groups, with a relative neglect of protective factors or groups which enjoy better mental health [9]. I therefore additionally intend this thesis to go some way towards correcting these more general research gaps.

1.2 PhD Aims

The purpose of this thesis is to describe and explain the apparent Indian mental health advantage in B-CAMHS. I shall do so by:

1) Characterising in detail the nature of the Indian mental health advantage.

2) Investigating how far the apparent Indian mental health advantage in B-CAMHS reflects a real ethnic difference in the prevalence of mental health problems, and how far it reflects inappropriate or biased measurement of mental health.

3) Investigating whether any real Indian advantage can be explained by the child, family, school and area characteristics measured in the surveys.

In addressing these aims, I focus upon comparing Indian and White children. This is because White children comprise the large majority of the British general population. They are therefore the ethnic group about whom most is known already, which makes them a useful comparison. Whites are also the only ethnic group larger than Indians, meaning that this is the best-powered ethnic contrast possible. The methodology developed form making this contrast may, however, lay the groundwork for future more extensive ethnic comparisons in datasets with larger numbers of children from other ethnic groups.

1.3 Outline of the PhD

This Introduction is followed by a series of Chapters which discuss concepts, methodologies and previous literature relevant to this thesis. Chapter 2 introduces child mental health as a health outcome, and highlights the key challenges in classifying and measuring mental health problems and disorders. Chapter 3 discusses the use of ethnicity as an explanatory variable, and outlines the migration history and current characteristics of the main minority ethnic groups in Britain. Finally, Chapter 4 presents a systematic literature review of ethnic differences in child mental health in Britain.

This provides the background for my quantitative analyses of B-CAMHS. Chapter 5 describes the survey methodologies and mental health measures used in B-CAMHS. Chapter 6 outlines some key analytic decisions which I have taken, and presents evidence on representativeness and participation rates in B-CAMHS. I then turn to my three PhD aims. Chapter 7 addresses the first aim through a detailed comparison of the mental health of Indian and White children for all available mental health outcomes. Chapter 8 addresses the second aim by examining issues of measurement. Specifically it evaluates which mental health outcomes show most evidence of cross-cultural validity and exploring other potential biases in the interview process. These analyses in turn form the basis for my choice of primary outcomes when addressing my third aim. Chapter 9 introduces these substantive analyses by describing the potential explanatory variables collected in B-CAMHS, and presenting a conceptual model for how these relate to ethnicity. This informs the univariable and multivariable analyses presented in Chapter 10 and Chapter 11, which seek to explain the Indian advantage in terms of the child, family, area and school characteristics of Indian children.

I conclude this thesis by summarising my findings and drawing conclusions. I then discuss the limitations of the analyses presented in this thesis, and highlight what information from future studies is most likely to make a valuable contribution to our understanding.

The main body of the thesis (Part 1) is followed by three Appendices (Part 2). Appendix 1 provides details of all statistical methods used or reported throughout this thesis; Appendix 2 provides information and analyses which supplement the research presented in the main body of the thesis; and Appendix 3 includes copies of all published papers arising to date from this PhD. Part 2 also contains the references from both the main body of this thesis and the appendices.

Chapter 2 Child mental health and child mental health problems

2.1 Introduction

2.1.1 Mental health as a continuum

The World Health Organisation (WHO) describes mental health as "a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community" [16, p.1]. Mental health, like physical health, is thereby conceptualised as more than the absence of disease. Rather it both has intrinsic value as a state of well-being and furthermore is a resource for achieving other goals [17]. These points apply with equal force to child mental health, although fulfilling one's potential, engaging with others and achieving personal and social goals will clearly take different forms in children compared to adults.

The term 'mental health' is, however, ambiguous. The WHO uses it above to refer specifically to good mental health, but it can also be used to encompass the whole mental health spectrum. To confuse matters further, in practice the vast majority of 'mental health research' focuses on mental health *problems*. Likewise the stigma surrounding mental health problems [18] has encouraged the euphemistic use of the term 'mental health' when providing psychiatric treatments and services [19].

In this thesis I use the term child mental health to describe the full of spectrum emotional and behavioural functioning. Within this, I define and differentiate its positive and negative manifestations as outlined in Box 2.1 and presented schematically in Figure 2.1. My underlying theoretical model is therefore of mental health as a continuum, with variation between individuals being observed across the full range and not simply between the 'mentally well' and the 'mentally disordered'. Focusing on mental disorders as an outcome may often be justified on the grounds that this marks a group with particularly severe and impairing problems, and also the group which is most likely to receive psychiatric treatment. Nevertheless, simply counting the number of people with a disorder may tell one less about the distribution of mental health in a population than measuring mental health as a continuous trait. Moreover, given that the population mean predicts the proportion of deviants for many health outcomes [20] including adult depression [21], understanding the determinants of mean mental health may enhance one's understanding of the prevalence of disorder. I therefore believe that continuous measures of mental health are important complements to binary measures of disorder, and this informs my analysis strategy throughout this thesis.

Box 2.1: Mental health terms used in this PhD

Mental health: Any aspect of behavioural, emotional or psychological functioning, including both positive and negative health experiences.

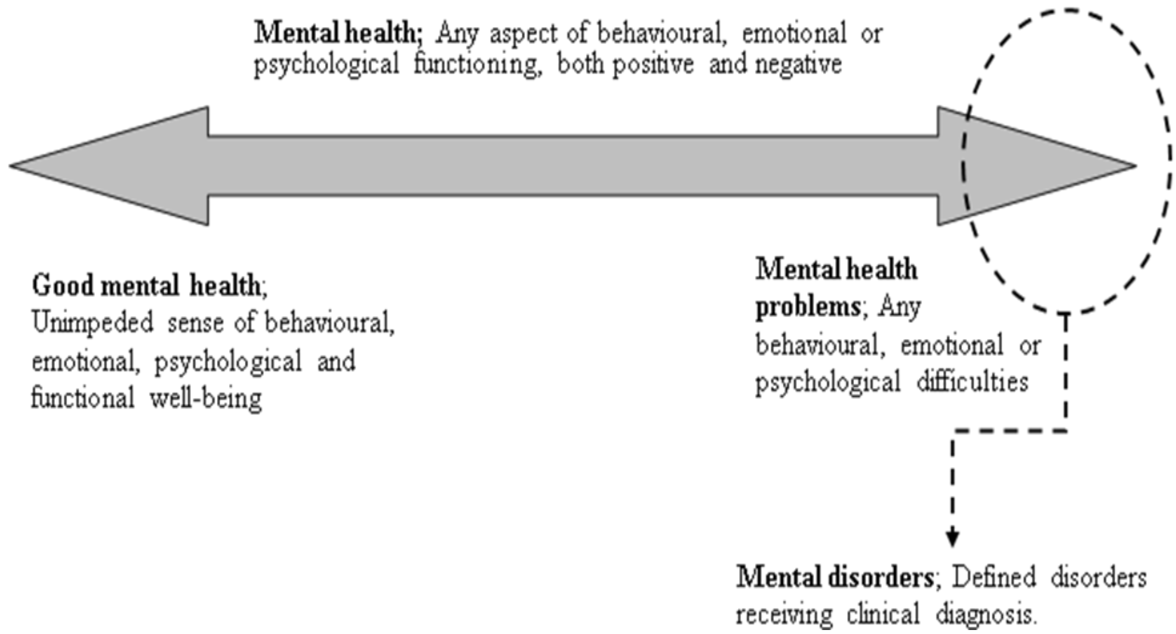
Good mental health: An unimpeded sense of behavioural, emotional and psychological well-being, in which the child experiences subjective well-being (life satisfaction), can realise their potential, cope with the normal stresses of life, and function productively in their relationships and activities (drawing upon [16, 22]).

Mental health problems: Any behavioural, emotional or psychological difficulties which reduce the subjective well-being of the child and/or impair the child's functioning in their everyday life (drawing upon [22]). Mental health problems therefore include mental disorders (see below), but also subclinical states and constellations of symptoms that do not fit standard diagnostic criteria.

Mental disorders: Defined disorders receiving clinical diagnosis. To qualify as a disorder requires that particular symptoms be present *and* have a substantial negative impact on the child, causing them distress and/or impairment in their everyday life. Using impact criteria means that any mental disorder should, by definition, also be a mental health problem.

Mental disorders are therefore a binary category, while good mental health and mental health problems are dimensional constructs (although both can be categorised if required).

Figure 2.1: Mental health as a continuum



2.1.2 A focus in this PhD upon child mental health problems

Conceptualising good mental health as more than the absence of disease implies that it deserves attention in its own right. For adults, the past two decades have seen a renewed emphasis on promoting good mental health [23-24] and an emerging science of adult well-being [25-26]. This has already confirmed the importance of good mental health for quality of life, physical health and social productivity [27-28]. It has also demonstrated that the absence of mental disorder is not synonymous with positive well-being, and that good mental health cannot be conceptualised, measured or explained simply as the inverse of mental health problems [29-31].

There has not yet been any comparable research program of good mental health in children [26]. A few large-scale research studies have investigated related constructs such as health-related quality of life [32] or the well-being of children as broadly defined [33], but the vast majority of child psychiatric epidemiology continues to focus upon mental health problems. This includes this thesis, which starts with the observation of a comparatively low prevalence of disorders in Indian children. This may, I hope, generate insights into *protective factors* against mental health problems. It does not, however, allow direct examination of good mental health and the factors which promote it.

My focus in this thesis is therefore on child mental health problems. As defined in Box 2.1, these include not only mental disorders, at the extreme negative end of the spectrum, but also sub-clinical states. This broader perspective follows from my theoretical model of mental health as a continuum. It also reflects my belief that all states and experiences which reduce child well-being are important problems, even if they do not meet the criteria for mental disorder.

2.2 The common child mental health problems

In this PhD, I focus upon the common child mental health problems, namely emotional, behavioural and hyperactivity problems. Later in this Chapter, I review critically the classification and measurement of these mental health problems, with a particular focus on the challenges of using them for cross-cultural comparisons. By way of orientation, I first summarise their key characteristics, including core symptoms, prevalence, long-term effects and association with child, family and area characteristics. Most of the studies which form the basis for this overview come from high income countries, especially the UK and US. In Section 2.3.3 of this Chapter I discuss how far these findings are replicated in low- and middle-income settings, and in Section 5.4, Chapter 5 I review the specific findings published to date from B-CAMHS.

2.2.1 Characteristics of the common child mental health problems

Emotional, behavioural and hyperactivity problems are the three broad domains of the most common child mental health problems. Their main clinical features are as follows [34]:

- **Emotional problems:** These are problems related to fears and phobias; worries and anxiety; depression, low mood and loss of interest; and obsessions.
- **Behavioural problems:** In these, children show a persistent failure to control their behaviour appropriately within socially defined rules. This is manifested in defiance, oppositionality, aggressiveness and/or antisocial behaviour.
- **Hyperactivity:** The core symptoms of hyperactivity are restlessness and fidgety behaviour; inattention and inability to concentrate; and impulsiveness.

As for ‘mental health’ more generally, these emotional, behavioural and hyperactivity domains all capture a mental health spectrum from very positive to very negative (see e.g. [35] on hyperactivity). At the negative extreme, each domain of problems includes a number of mental disorders, as listed in Box 2.2. Behavioural and hyperactivity problems are also frequently combined to form a single category, which is then contrasted with emotional problems. At different points in this thesis I use both the three-way and the two-way division of the common child mental health problems. When doing so I follow the convention of referring to the former as ‘emotional’, ‘behavioural’ and ‘hyperactivity’ problems and the latter as ‘internalising’ and ‘externalising’ problems.

In addition to the disorders listed in Box 2.2, there are also a large number of ‘less common disorders’ which can affect children. These are a group of rare to very rare conditions, the main disorders being pervasive developmental disorders (including autism/autistic spectrum disorders), psychotic disorders, tic disorder and eating disorders. The heterogeneity of these conditions means, however, that it makes little sense to talk of ‘less common mental health problems’ as a single meaningful entity. In conjunction with inadequate power, this consideration means that I do not make much use of the less common mental disorders in this PhD.

Box 2.2: Mental disorders in the three broad domains of child mental health problems[2]

Emotional disorders (also called internalising disorders):
Anxiety disorders

- Separation anxiety
- Specific phobia
- Social phobia
- Panic attacks
- Agoraphobia
- Post Traumatic Stress Disorder (PTSD)
- Obsessive-Compulsive Disorder (OCD)
- Generalised Anxiety Disorder (GAD)
- Anxiety disorder, Not Otherwise Specified

Depression

- Depressive episode
- Depressive episode, Not Otherwise Specified

Behavioural disorders (grouped with hyperactivity disorders to give externalising disorders)

- Oppositional defiant disorder (ODD)
- Conduct disorder
- Disruptive Behavior Disorder, Not Otherwise Specified.

Hyperactivity disorders (grouped with behavioural disorders to give externalising disorders):

- Attention Deficit Hyperactive Disorder (ADHD) Combined Type
- ADHD Inattentive Type
- ADHD Hyperactive-Impulsive Type
- Hyperactivity disorder, Not Otherwise Specified

2.2.2 Prevalence, comorbidity and time trends

Prevalence

Epidemiological studies of child mental health problems vary greatly in their methodology, including in the informant approached, the questionnaire or interview administered, and the diagnostic criteria applied. Comparing prevalence estimates across studies is therefore difficult, with a relatively wide range of estimates even from apparently similar settings.

Nevertheless, several decades of research have made it clear that child mental health problems are common: a review of 49 large, population-based surveys from around the world found that most reported 10-20% prevalence for any disorder, giving a mean of 12.9% [4]. Emotional disorders, particularly anxiety disorders, are often the most common disorders with a prevalence of 5-10% [36-37]. Behavioural disorders are typically somewhat less common with a prevalence of 3-5% [38-39]. Hyperactivity disorders are

[2] Diagnoses presented are those recognised under the DSM-IV classification system (see Box 2.3, p.43). The disorders recognised by ICD-10, the main alternative system, are very similar.

less common still, with 0.5-4% prevalence for ADHD and 0.5-2% prevalence for the more restricted hyperkinetic disorder [40]. The prevalence of most less common disorders is typically well under 1% [4].

Comorbidity

Epidemiological studies have also confirmed clinical observations of frequent comorbidity between child mental health problems. Comorbidity is highest between behavioural and hyperactivity problems, with around half of all children with a hyperactivity disorder also receiving a diagnosis of a behavioural disorder [40-42]. The reasons for this comorbidity are unclear, and may include a mixture of common underlying pathology, shared risk factors (either genetic or environmental) and hyperactivity increasing vulnerability to subsequent development of behavioural problems (e.g. by exposing the child to frequent criticism).

Comorbidity between behavioural and emotional problems is also high. For example, a review of 21 population-based studies using comparable diagnostic criteria estimated odds ratios for behavioural disorders of 6.6 for children with depression and 3.1 for children with an anxiety disorder [43]. Again, the reasons for comorbidity are unclear. Some of the association may reflect a direct causal role from one disorder to the other – for example, a behavioural disorder may cause difficulties with friends and family that then precipitates emotional problems [44]. There is, however, some evidence that the majority of the comorbidity reflects shared environmental risk factors [45] and shared genetic liability [46]. Shared genetic liability may, in turn, be mediated by common temperamental risk factors such as high mood lability [47].

Time trends

As with all comparisons of prevalence, methodological variation between studies complicates the examination of time trends. There is, however, evidence that over the past few decades child mental health has worsened in Britain, the US and several other high income countries [6]. Particularly notable has been a rise in behavioural problems. For example, Collishaw *et al.* [7] use comparable measures to compare three representative samples of British 16 year olds born in 1958, 1970 and 1983[3] There was a consistent rise across the three samples in parent-reported behavioural problems. Moreover, the long-term implications of childhood behavioural problems had grown no less severe over time, suggesting that the increase was real rather than a reporting bias. The reasons for the trend are, however, largely unknown: problems increased across all family types, family sizes, and socio-economic groups, and none of these factors seemed to play a large role in explaining the increase.

Among the other common mental health problems, there was some evidence of a modest increase in emotional problems in the most recent sample but no evidence of a change for hyperactivity problems. Both emotional and particularly hyperactivity problems did, however, show increasing comorbidity with behavioural problems over time. The reasons for this are not known, but it may be that children with emotional or hyperactivity problems are particularly vulnerable to whatever risk factors are causing the rise in behavioural problems more generally.

2.2.3 Long-term effects

Child mental health problems and disorders are therefore common, frequently comorbid, and may be increasing in prevalence. This is a major cause for concern because of their implications for both short-term and long-term well-being. As previously described in Box 2.1, a necessary criterion for diagnosis is that symptoms are causing the child acute distress and/or marked impairment. Child mental health problems and disorders also predict many other adverse outcomes for the child and for those who care for them. This is emphasised in Section 2.2.4, which stresses that child mental health problems are likely to be one cause of factors such as poor physical health, academic difficulties and poor parent mental health.

Moreover, these negative effects are frequently not short-lived. In the medium term, B-CAMHS99 found that at a group level the distress and impairment experienced by children with a disorder was unchanged over three years [48]. In the longer term, emotional and

[3] These samples came from the National Child Development Study (1958 birth cohort), the British Cohort Study (1970 birth cohort) and the British Child and Adolescent Mental Health Survey 1999.

behavioural disorders in children frequently persist into adult life [49], and there is growing evidence that the same applies to hyperactivity [50-51].

This is serious for many reasons. First, suicide is a leading cause of death for adolescents and young adults [52-54], and mental disorder plays a central role in predicting suicidal behaviours [55-56]. Furthermore, childhood mental health problems also predict negative outcomes in other domains of adult life. In late teenage and early adult life, children with hyperactivity problems are at higher risk of academic underachievement [57-58], while children with behavioural problems are at higher risk of criminal behaviour and substance abuse [58-60]. Childhood behavioural problems also predict adverse events across a wide range of domains in later adulthood. For example, British 16 year olds with 'high' scores for behavioural problems had a substantially greater likelihood of experiencing unemployment, homelessness, teenage parenthood and alcohol problems by the age of 30 [7]. Emotional disorders in the teenage years likewise increase the risk of educational underachievement, unemployment and poor health, with the effect on the latter being particularly marked [61-64]. Indeed, even questions on individual emotional symptoms in early childhood have recently been shown to predict permanent sickness or disability in mid-adulthood [65].

2.2.4 Correlates of child mental health problems

The previous section illustrated that emotional, behavioural and hyperactivity problems show some specificity in the type of adverse outcomes which they predict [55]. Some specificity is likewise observed in their cross-sectional associations with child, family, area and school characteristics. Below I summarise the key correlates of child mental health problems. This provides a context for my discussion in Section 5.4 Chapter 5 of the B-CAMHS findings, and for my subsequent investigation into which characteristics may explain the Indian mental health advantage. This section also introduces one key challenge in studying the determinants of child mental health, namely that for many 'risk factors' the direction of causality with child mental health is unclear. I return to this challenge in Section 11.1.2, Chapter 11.

In this section I do not present evidence on the association between child mental health and ethnicity, as this forms the focus of my systematic review in Chapter 4.

Child characteristics

Age and sex

In almost all clinical and epidemiological studies ever conducted, boys show more behavioural and hyperactivity problems while girls (particularly older girls) show more emotional problems [4, 36-40, 42]. This has been replicated across countries and in studies over the past 50 years, although for some disorders such as conduct disorder there is evidence that the gender difference has narrowed in recent decades [39]. Another highly consistent pattern across time and space is the decline of hyperactivity after age 6-9 years [40] and a sharp increase in depressive disorders in adolescence [37]. This often includes an age-gender interaction such that the female disadvantage for emotional problems is larger in the teenage years than at younger ages.

Physical health and development

Non-specific somatic complaints such as headaches and stomach aches are a common symptom of anxiety or depression in children. It is therefore unsurprising that poor self-reported general health is particularly strongly associated with emotional problems. Yet this cannot be the whole explanation, as child mental health problems – particularly emotional problems – are also associated with specific physical disorders such as diabetes. Around 30-40% of children with severe paediatric illness have a comorbid child mental disorder. A modest increased risk is also associated with less severe conditions like asthma [66].

Most specific physical disorders show the same pattern as self-reported general health in being more strongly associated with emotional problems than with behavioural or hyperactive problems. Brain damage is one notable exception, having a particularly strongly association with hyperactivity [67]. In the Isle of Wight studies of the 1970s, for example, children with brain damage were seven times more likely to suffer from any mental disorder and 90 times more likely to suffer from hyperactivity [68]. This is consistent with the strong association between hyperactivity and many other markers of developmental problems or delays, including language retardation or speech and motor impairments [40].

Cognitive function and academic abilities

The neurocognitive aspect of hyperactivity problems is also indicated by a strong association between hyperactivity and poor cognitive functioning, including low IQ and poor progress in school [40]. This association remains even after excluding children with brain damage, developmental delay or learning disorder. There also appears to be a dose response relationship such that those with more severe hyperactivity have greater cognitive impairments [35, 69]. There is some evidence that the two conditions may share underlying genetic risk factors [70-71].

Yet while the association is particularly strong for hyperactivity, low IQ and/or poor school performance are also associated with behavioural and emotional problems. For behavioural problems the association with low cognitive function applies particularly to poor verbal and planning skills, and may reflect shared underlying traits (e.g. impulsivity) [39]. The stress of having below average ability in school also seems important, however, as indicated by the poorer mental health of children who are the youngest in their school year [72]. Yet while this provides evidence of at least some forward causal role of academic performance, reverse causality is also highly plausible. The relative strengths of the effects in each direction are largely unknown.

Parenting experiences

Various parenting behaviours predict future mental health. At the extreme are strong negative effects from a range of serious parental failures, including child abuse [73] and routine foster care [74-75]. For example, recent surveys of looked-after children in England found rates of disorder of almost 50%, compared to the general population average of 10% [75].

Several less extreme parenting behaviours are also associated with child mental health problems. These include hostile parenting, poor supervision, or a parenting style which is either punitive/authoritarian or indulgent/indifferent [reviewed in 39, 76]. In addition, a large number of studies have indicated an association between child mental health problems and parental use of punishment, including harsh physical punishment [39, 77-79].

The direction of causality for these associations is unclear; for example, authoritarian parenting or the use of punishment may sometimes be a response to difficult child behaviour rather than a cause of it. Nevertheless, the positive effect upon child behaviour of interventions to improve parenting skills [e.g. 80] indicates the likelihood of some forward causal component.

Stressful life events

A substantial literature indicates an association between accidents/unintentional injuries and externalising problems, with the link with hyperactivity being particularly well-described [81-84]. For severe head injuries, a biologically plausible mechanism exists whereby the accident may cause the development of symptoms of hyperactivity [85]. Yet most accidents and injuries are not of this sort, and the association between externalising problems and other forms of accidents (e.g. burns) seems likely to reflect children with externalising problems behaving in more dangerous ways. Moreover, even for head injuries, relatively convincing evidence from the UK suggests that the association between early head injury and later diagnosis of ADHD is not causal. Rather early injury may be a marker for pre-clinical symptoms which predict subsequent diagnosis [86].

The association between accidents/injuries and externalising problems may therefore primarily reflect a reverse causal relationship. By contrast, there is better evidence of a forward causal relationship between being in a serious accident and the development of emotional (particularly anxiety) disorders [87]. Other acute stressful life events for which there is evidence of an effect include witnessing a trauma; family crises such as repossession of the home; and the loss of a parent, sibling or friend through death or divorce [87-88]. Inter-personal violence is also important as both an acute and a chronic stressful life event. This not only includes violence within the family (e.g. harsh physical punishment) but also violence by peers, teachers or during wars [78, 87].

Friendships and peer relations

Many aspects of peer relations have important effects on child mental health. Children show large individual differences in their resilience to stressful life events such as those described above [87], and difficulties with peers is one factor making children more vulnerable [89-90]. Difficulties with peers are also predictive of future emotional disorders [89, 91] and of poorer prognosis in children with existing emotional disorders [92].

Yet while an absence of friends is a risk factor for emotional problems, the effect of peer groups may not always be positive. In particular, participation in delinquent peer groups is strongly associated with antisocial behaviour [39]. Some of this undoubtedly reflects selection effects, whereby defiant or disruptive children are drawn to deviant peer groups. There is, however, also evidence of that deviant peer groups play an independent role by reinforcing and perpetuating antisocial behaviours.

Substance use

It is well-established that externalising problems and disorders in early or middle childhood predict substance use, abuse and dependence in adolescence and young adulthood. With very few exceptions [93], this has been shown for tobacco [64, 94-99], cannabis [64, 97, 100-105] and other illicit drug use [97, 106]. This has also been shown for alcohol [97, 100, 102, 106-108], although the effect is often weaker and not always observed for alcohol use (as opposed to abuse) [61, 109]. A growing body of evidence indicates that these associations are mainly driven by behavioral problems – hyperactivity has little or no independent effect [97, 101, 107, 110]. For internalising problems the evidence is less clear. Interpreting many studies is complicated by a failure to adjust for possible comorbidity with externalising problems [111-114]. Among studies which do control for comorbidity, some report independent effects upon substance use or dependence [93, 101, 115] while others do not or report only weak or inconsistent effects [98, 103-104, 116-117].

As for the effect of substance use upon mental health, there is consistent evidence that smoking cigarettes or cannabis in adolescence predicts anxiety [116, 118-122], depression [97, 118, 123-126] and perhaps behavioral problems [119, 127] in early adult life. Fewer studies, however, examine mental health outcomes in the teenage years. These produce mixed results [93, 98, 127-129], including negative or inconsistent findings [98, 127-128].

Family composition

Family type

Many cross-sectional studies and some longitudinal studies demonstrate poorer mental health outcomes in children living in lone-parent or stepfamilies, as compared to children living with both biological parents [130]. This evidence includes data from the two large British birth cohorts of 1958 [131-132] and 1970 [133]. These studies were, however, unable to examine in detail the relative contribution of family type *per se* as distinct from associated factors such as reduced family income. The 1997 Health Survey for England suggested that socio-economic adversity explained the mental health disadvantage of children in lone-parent families but not in stepfamilies [130].

Family size

There is some evidence that children with many siblings suffer more behavioural disorders, with the number of older brothers seeming particularly important [39]. This increased risk may be largely mediated by ineffective parenting and family conflict.

Family stress

Family conflict and family functioning

Parental conflict and poor family functioning is associated with behavioural problems, emotional problems and poor social functioning [134-135]. There is some evidence that these effects remain after adjustment for age, gender and a range of other individual and family factors [136-137].

Many of these studies are cross-sectional, meaning that the direction of causality is unclear. Some bidirectionality is plausible, however, given the potential for mental health symptoms (particularly externalising symptoms) to elicit self-reinforcing cycles of negative interactions with family members [138].

Parent mental health

Child mental health problems are associated with a range of parent mental health problems, including parental alcoholism, antisocial behavioural, schizophrenia and depression [139]. Much of this research focuses on maternal mental health. For example longitudinal studies show that maternal depression predicts poorer outcomes for child development and mental health in infancy [140], pre-school [141], early childhood [142] and later childhood [143]. More recently, paternal depression has also been shown to have negative effects through early and middle childhood, independent of maternal depression [144-145]. Depression in either parent may also increase vulnerability to other risk factors such as stressful life events [146].

The highest correlation is observed between child and parent disorders of the same type, although it is also observed across disorder types [139]. For example, the children of adults with depression are at greater risk of mental health problems in general and depression in particular. To a large extent this may reflect shared genetic liability, but it may also reflect non-genetic learning of (for example) particular attributional styles. Certainly there is a strong correlation in the timing of depressive episodes among parents and children who have a history of depression [147], suggesting that the one is triggering the other. Here, as for the relationship between child and parent mental health more generally, bidirectional causation is likely [148].

Family socio-economic position

Low socio-economic position (SEP) is associated with poorer outcomes for children's cognitive development, physical health and social well-being [149]. A large number of cross-sectional and longitudinal studies have indicated that low family SEP (variously operationalised in terms of occupation, income or parent education) is also associated with higher rates of emotional and behavioural problems. This is true in both childhood and adolescence, and the relationship shows a dose response relationship across the whole SEP range [42, 139]. For hyperactivity the evidence is more mixed; some studies show an association with SEP but many do not [35, 40].

In most cases the effects of low SEP seem to be mediated via factors such as parent mental health and family conflict [39]. This does not diminish the importance of socio-economic disadvantage as an upstream determinant, however, especially given some evidence that poverty relief can improve child mental health [150].

Area characteristics

Area deprivation

Area deprivation is associated with child mental health problems, particularly behavioural disorders [139], with effects persisting even after controlling for family SEP [151-152]. There is some evidence that these effects relate primarily to the characteristics of the other people who live in such areas rather than to properties of the areas *per se*, but relatively little research has been done on this topic [39].

Area ethnic density

Ethnic density refers to the proportion of residents from a particular ethnic group. That the density of one's own ethnic group may influence adult mental health has some support from studies in the UK [153-155] and in other countries [reviewed in 156]. Little is known about ethnic density and child mental health, although one UK study reports that South Asian children had the best mental health in areas of moderate ethnic density [157].

Most studies hypothesise a protective effect for ethnic density. High ethnic density is suggested to promote good mental health through multiple mechanisms, including facilitating social cohesion and integration, providing social support and serving as a buffer against racism [154, 158]. Many studies fail, however, to deal adequately with the important confounder of area deprivation; because many areas of high ethnic density are also relatively deprived then this could lead to any ethnic density effect being masked [156]. This may partly explain why the observed relationship between ethnic density and child or adult mental health has been inconsistent, ranging from a linear protective effect [154] to a U-shaped relationship [157] to an inverted U-shaped relationship [155]. In conjunction with the limited evidence base, these inconsistencies make it impossible to draw conclusions about the effects of ethnic density upon child mental health.

School characteristics

There has been little research into the effect of schools on child mental health. What evidence exists, however, indicates a potentially important role both for the composition of the school's pupils (e.g. the proportion from socio-economically disadvantaged families) and for the school's own attributes (e.g. its ethos) [39].

2.2.5 The importance of genes

While genetic analyses do not feature in this PhD, there is no doubt that genes play a major role in determining a child's vulnerability to mental health problems. There is substantial evidence of the high heritability of many mental disorders or mental health-relevant traits [159]. This applies to both externalising and internalising problems, although heritability is higher for externalising problems and particularly for hyperactivity [160].

Many of the continuities between child mental health problems and adverse adult outcomes reviewed in Section 2.2.3 may to some extent be mediated by shared genes [159]. Many of the 'environmental' effects discussed above may also have some genetic component. This is illustrated by gene-environment correlations in which particular alleles increase the risk of certain stressful life events [161]. The substantial heritability for many risk factors such as divorce, poor parenting or substance use [162-163] likewise indicates that some apparent environmental effects may be due to confounding by shared genes. Disentangling these effects requires sophisticated study designs such as adoption studies or, more recently, using in-vitro fertilisation pregnancies [164].

Yet such attempts at disentangling genetic and environmental effects may not always be the most informative approach. Instead, a growing body of evidence suggests that often it may be the *interaction* between particular alleles and particular environmental exposures which is key in determining individual risk. Examples include genetic interactions with life events in causing depression in girls [146] and with child abuse in causing antisocial behaviour in boys [165]. This last example is also unusual in that the gene in question has been identified (Monoamine Oxidase A), thereby representing one of the few convincing examples to date of a common allele with a substantial mental health effect.

2.2.6 Summary and conclusion

To summarise, emotional, behavioural and hyperactivity problems and disorders differ in their symptoms, patterns of comorbidity and key correlates. All, however, cause substantial distress and impairment. It is therefore a matter of great concern that 10-20% of children suffer from a mental disorder at any one time; that many of these children have more than one disorder; and that the prevalence of behavioural and perhaps emotional problems seems to have increased in several countries including Britain. Child mental health problems also predict reduced well-being in later life, including poor mental health, poor physical health, and constrained life opportunities.

Yet despite their major public health importance, studying child mental health problems raises several important challenges. One is the likelihood of a bidirectional causal relationship with many child and family factors, which complicates the interpretation of cross-sectional associations. There are also other features of the common child mental health problems and disorders which create particular challenges for definition and assessment. The challenges cut across all psychiatric epidemiology, but are arguably particularly important in cross-cultural comparisons. These issues are therefore central to this thesis, and form the subject of the remainder of this Chapter.

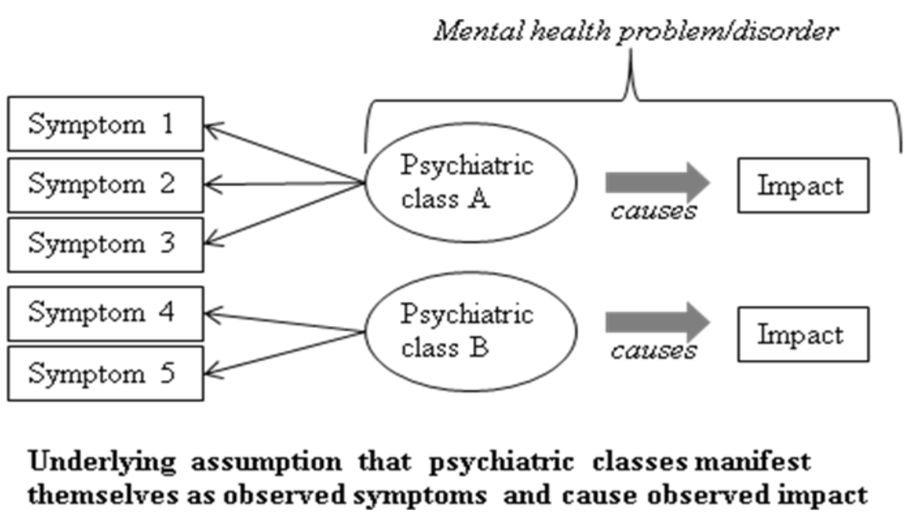
2.3 Classification of child mental health problems

2.3.1 The phenomenological approach to classification and measurement

Child psychiatry is a long way from being able to identify or quantify mental health problems on the basis of a detailed understanding of the underlying pathophysiology. Instead the dominant approach has since the 1960s been phenomenological, with mental health problems classified through detailed descriptions of observed symptoms [166]. Psychiatric classes therefore do not describe underlying diseases which explain *why* a child behaves a certain way [167]. Rather, as Figure 2.2 schematically represents, they are defined as groups of symptoms which frequently co-occur. This applies both to narrow classes such as 'separation anxiety' and to broader classes such as 'emotional problems'.

More recently, substantial negative impact has been added as a necessary criterion for considering a symptom cluster to represent a mental disorder. This is important in preventing unnecessary and unhelpful diagnoses for children who experience no negative consequences from their symptoms. This in turn avoids misleadingly high population prevalence estimates of up to 50% [e.g. 168, 169]. Conversely, impact criteria may also facilitate the appropriate assignment of Not Otherwise Specified diagnoses (Box 2.3, p.43) to children with important mental health problems and who might benefit from services.

Figure 2.2: Phenomenological classification of mental health problems and disorders



The phenomenological approach to classification is reflected in current approaches to mental health assessment. Dimensional questionnaires typically assess the extent to which symptoms from a particular class are present, and sum these to give overall scores. As described in Box 2.3, the dominant diagnostic classification systems likewise operationalise disorders in terms of symptoms and impact, together with some additional criteria about duration and onset and additional rules about comorbidity.

Box 2.3: Diagnostic classification systems: DSM-IV and ICD-10

At present, the two dominant, internationally-recognised systems of diagnostic research criteria are the Diagnostic and Statistical Manual of Mental Disorders, 4th edition [DSM-IV; 170] and the current International Statistical Classification of Diseases [ICD-10; 171]. These systems are very similar in their approaches to classification and in most details of how they operationalise particular disorders.

Similarities in general approaches to classification:

- **Multiaxial classification**: DSM-IV and ICD-10 are multiaxial systems, meaning that they recognise that children simultaneously occupy a position on multiple different axes of functioning. For instance, DSM-IV's Axis I records clinical disorders, Axis II records personality disorders and mental retardation, Axis III records general (physical) medical conditions, Axis IV records psychosocial problems and Axis V makes a global assessment of functioning [170]. In B-CAMHS only Axis I assessments were made – i.e. clinical diagnoses of mental disorders. These form the focus of my discussion below and throughout this PhD.
- **Obligatory inclusion and exclusion criteria of symptoms.** Most diagnoses for mental disorders have symptom inclusion criteria, such that they require the presence of certain symptoms or a minimum number from a list (e.g. at least four out of 10). Some diagnoses also have exclusion criteria, where a diagnosis *cannot* be made if a certain symptom is present.
- **Impact criteria.** A necessary criterion for diagnosis is that the symptoms are causing the child acute distress and/or marked impairment in everyday life.
- **Duration and onset criteria.** Some disorders require symptoms to be present over a minimum time period (e.g. the past two weeks) and/or to have started before a

maximum age (e.g. onset before age three).

- **Not Otherwise Specified diagnoses.** These can be assigned to children with symptoms causing them substantial distress or impairment, but who do not meet the symptom-based criteria for an operationalised disorder [34, 172].
- **Comorbidity.** Children can simultaneously receive more than one diagnosis, although DSM-IV and ICD-10 differ somewhat in their extensions (e.g. mixed categories) and restrictions (e.g. hierarchies of disorders) [167].

Similarities and differences in criteria for specific disorders

- DSM-IV and ICD-10 are generally very similar in their criteria for specific disorders [167].
- The one major difference is between Attention Deficit Hyperactivity Disorder (ADHD) in DSM-IV and hyperkinetic disorder in ICD-10. While the symptoms for the two conditions are near-identical, ICD-10 requires more symptoms, greater pervasiveness of symptoms, and has stricter exclusion/comorbidity criteria. Approximately 30% of children who meet the diagnostic criteria for ADHD therefore do not meet the criteria for hyperkinetic disorder.

2.3.2 Implications for establishing the validity of psychiatric classes

Building a case for validity

The development of agreed diagnostic schemes and clearly-defined criteria for child mental disorders has played a crucial role in increasing inter-practitioner reliability and inter-study comparability [167]. Nevertheless, basing classification upon observed symptoms rather than upon underlying pathological mechanisms raises important difficulties for establishing the validity of psychiatric classes. Not all children have all symptoms, creating the problem of how much variation is permissible within the 'same' class. Conversely, many children have symptoms from across two or more nominally different classes. This creates controversy as to whether a classification system is adequate or whether it should be modified by, for example, combining classes together or proposing alternative categorisations.

It will rarely be possible to resolve these issues definitively. Instead one needs to build a case for the validity of a particular system by, at a minimum, showing that the proposed symptoms do cluster together and are associated with negative impact. When proposing new psychiatric classes, it is also necessary to *justify* those classes as making useful predictions about external factors such as differential prognosis or treatment response [167]. Some flexibility in operationalising classes is also common, as exemplified by criteria requiring a minimum numbers of symptoms from a list and by the existence of Not Otherwise Specified categories (see Box 2.3, p.43).

The need to re-establish validity in all new populations

The phenomenological approach means that the case for validity rests not upon generalisible understandings of disease mechanisms but upon observations of symptoms and impact in particular populations. There may therefore be no classification system which is uniquely best or universally applicable. For example, symptoms which fall into separate clusters in high risk populations may form a single cluster in low-risk populations. This seems to apply to adult emotional problems, with the distinction between depressive and anxiety symptoms being far more apparent in clinic than in community settings [173].

Symptom-clusters may also differ across social or cultural populations such that classes are meaningful in some populations but not others. This possibility forms the basis for the relativist critique of the universalist approach to cross-cultural psychiatry. The universalist position stresses the shared features of mental health problems and disorders in different populations and seeks to study these using a single standard framework [78, 174-175]. This 'etic' approach underpins most psychiatric epidemiology and lies at the heart of diagnostic classification systems like DSM-IV [176]. Yet DSM-IV largely developed through European and North American input [177], and has been criticised for showing Euro-centric bias [178-180]. Most dimensional questionnaires and empirically-derived symptom structures were likewise developed in Western populations.

The relativist critique highlights the danger of assuming that these Western-derived classification systems will apply universally. One central tenet of the relativist position is the importance of considering 'emic' mental health phenomenology, namely that which is meaningful within any particular culture [181-182]. A second tenet is that illness experiences in different cultures may be so different as to represent genuinely separate conditions. Applying nominally etic classification systems therefore risks committing the 'category fallacy' of carving up the mental health landscape in a way which lacks face validity and coherence [14, 181, 183-184].

In thinking about category fallacies, I believe distinguishing face validity and coherence is crucial. 'Face validity' category fallacies correspond to the relativist concern for emic understandings, and occur whenever a particular construct does not map onto a locally meaningful category. This would certainly represent grounds for caution, but would not necessarily prevent meaningful cross-cultural comparisons or the useful application of insights about aetiology, prognosis or treatment. For example, in rural Uganda no local concept maps straightforwardly onto the English term 'depression'. This did not, however, prevent the DSM–IV criteria for major depressive disorder being used to identify adults who were successfully treated with interpersonal psychotherapy [185].[4] So long as similar constellations of symptoms exist in different populations, I believe it may be possible to apply etic classifications in a meaningful way. By contrast, 'coherence' category fallacies occur if symptoms show fundamentally different patterns of association in different populations. This form of category fallacy is the more serious threat, as it renders comparisons genuinely meaningless.

Finally, the centrality of negative impact to the definition of mental health problems or disorders means one must also remember that a given symptom-cluster may be a 'problem' in one population but not in another. For example, one might identify a population in which a comparatively high proportion of children displayed the core symptoms of hyperactivity (e.g. restlessness and short attention span), but where this did not cause them distress or impairment. It might certainly be of interest to study hyperactivity symptoms in that population, not least to investigate factors which were protective against negative

[4] This issue also applies to physical illnesses. For example, early on in the HIV/AIDS pandemic before a blood test was available, HIV/AIDS was diagnosed based on clinical symptoms and signs. Most of these were derived from observations in the USA, and HIV/AIDS certainly lacked face validity in many Sub-Saharan African settings which had no equivalent in their lexicon or nosology.

impact. In the absence of such impact, however, it would not be appropriate to consider hyperactivity as a highly prevalent mental health *problem* in that particular population.

Cross-cultural category fallacies as an instance of a more general challenge in psychiatric epidemiology

Criticism of the universalist approach as an unexamined default position has mainly occurred within the field of cross-cultural psychiatry. Yet demonstrating that proposed psychiatric classes are internally coherent and reduce well-being is crucial in *any* population. Likewise, establishing the comparability of the constructs under consideration is essential for meaningful comparison across any populations, including across time, space or social group. Moreover, there are no absolute criteria for deciding how much variation between populations is permissible before a category fallacy occurs. This has strong parallels with the difficulty of deciding how much variation between *children* is permissible within the 'same' disorder.

I therefore believe that the relativist critique is closely linked to a central challenge in all psychiatric epidemiology, namely the need to build a case for the validity of any classification system in any population to which it is applied. In the next section I discuss the evidence on this issue for the broad domains of common child mental health problems which I use in this thesis. In doing so, I first present evidence from the UK and similar settings and then discuss how far this is replicated in other cultures. Because little has been published regarding children from different ethnic groups in the UK, I instead draw upon the wider cross-cultural literature.

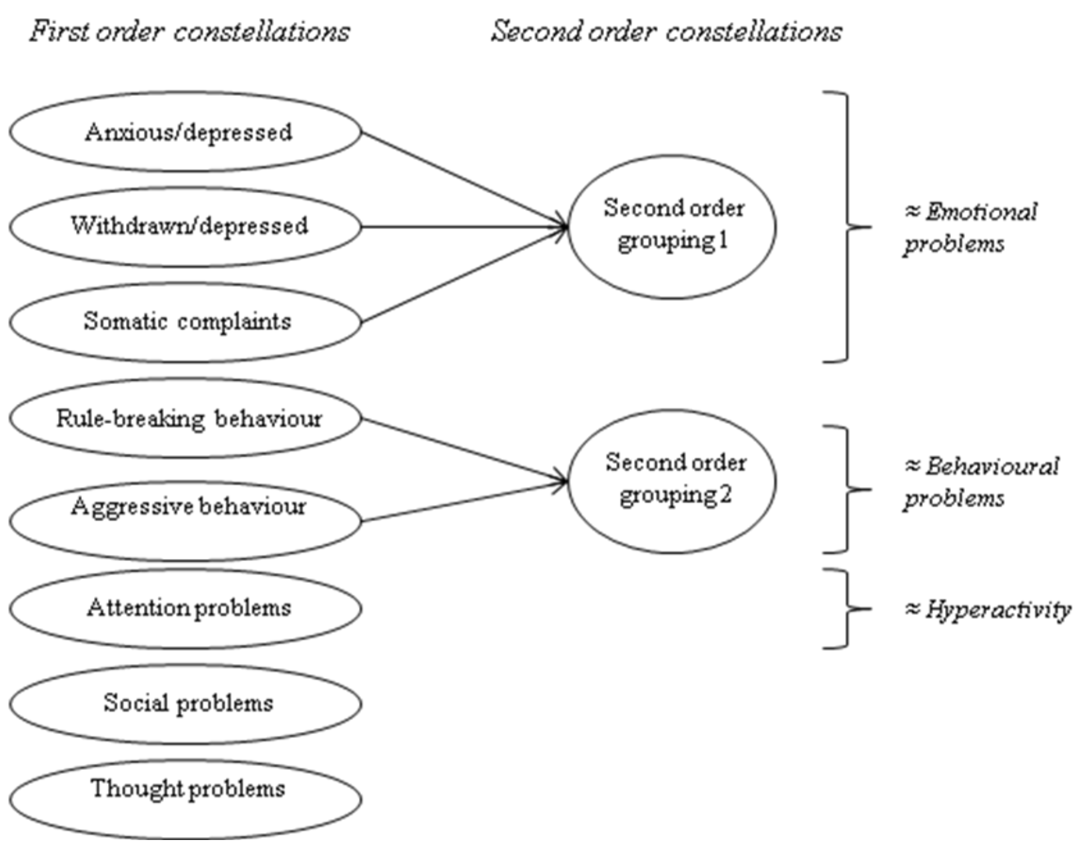
2.3.3 Validity of the common mental health problems and disorders as psychiatric classes

<u>Empirically derived symptom structures</u>

A 1946 study of 500 children in an American clinic provides an early example of attempting to validate psychiatric classes. The study used both factor analyses and associations with family correlates to argue for a distinction between over-inhibited behaviour (roughly corresponding to emotional problems) and two sorts of disruptive behaviour (socialised and unsocialised) [186].

Many subsequent studies have used similar techniques to derive and/or confirm symptom structures. This has been particularly evident for the Achenbach System of Empirically Based Assessment (ASEBA) for parents, teachers and children [187-191].[5] The ASEBA were initially developed in the 1960s and 1970s to identify symptom clusters ('syndromes') not included in the then-current DSM, and have since been subject to ongoing evaluation [42, 192]. Using factor analyses and other techniques on large clinic and community samples, the questionnaire authors empirically derived eight syndromes and two second-order groupings. These showed fairly good correspondence with the constructs emotional, behavioural and hyperactivity problems (see Figure 2.3).

Figure 2.3: Syndrome constellations identified on the ASEBA

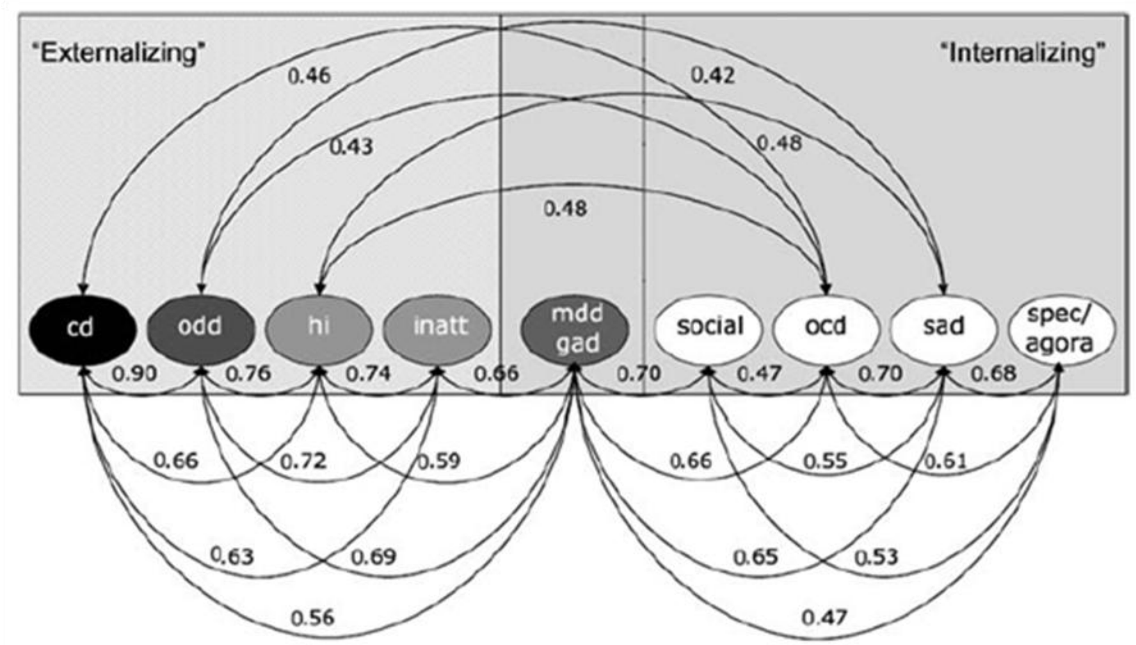


Similar findings have been reported for other widely used questionnaires, including the Rutter [68, 193] and the Strengths and Difficulties Questionnaire (SDQ) [194]. As reviewed in Section 5.3.2 Chapter 5, factor analyses universally support the distinction

[5] These are sometimes referred to by their separate names: the Child Behaviour Checklist (CBCL) for parents, the Teacher Report Form (TRF) for teachers and the Youth Self Report (YSR) for children.

between ‘internalising’ and ‘externalising’ symptoms in the SDQ, and often more detailed distinctions too. Further support comes from recent detailed analyses of the correlation structure of DSM-IV symptoms for common child mental disorders [195-196]. For example, Figure 2.4 demonstrates the observation of the expected second-order internalising and externalising factors in a representative sample of 4049 American children [195]. Yet Figure 2.4 also suggests some possible inadequacies of the DSM-IV system. In particular, symptoms of major depressive disorder and generalised anxiety disorder were indistinguishable in factor analyses, and their joint factor was highly correlated with both internalising and externalising factors.

Figure 2.4: Correlation among the latent factors in the best-fitting model for parent-reported DSM-IV symptoms



cd= conduct disorder, odd= oppositional defiant disorder, hi= hyperactivity–impulsivity, inatt = inattention, mdd = major depressive disorder, gad = generalised anxiety disorder, social = social phobia, ocd = obsessive-compulsive disorder, sad = separation anxiety disorder, spec = specific phobia; agora = agoraphobia.

All correlations >0.4 presented

Source: Lahey *et. al* [195, p.196]

Most studies of empirically derived symptom structures come from Western populations, but these methods have recently been applied more widely. For example, the ASEBA structure in Figure 2.3 has shown good fit to the data across 20-30 diverse countries, including several middle and low-income settings [42, 197-199]. The anticipated structure of the SDQ has likewise been demonstrated in many middle- and low-income settings including Brazil, Bangladesh, Pakistan and the Yemen [reviewed in 3 (Appendix D), 191, 200].

Etic comparisons of disorders

The above analyses are particularly informative because they take a ‘bottom up’ approach, starting with a diverse collection of symptoms and using empirical techniques to find symptom constellations. A number of reviews also evaluate how far existing nosologies (e.g. DSM-IV) fit child mental disorders from around the world. In general, these find that the characteristics of child mental disorders show broad and sometimes striking similarities [78, 174, 178, 201-205]. For example a review of WHO case studies in low- and middle-income countries (including India) concluded that there was relatively little difference across settings in the symptoms of particular disorders [205]. The review also concluded that culture-bound syndromes were very rare among children, and to my knowledge no culture-bound syndrome has been reported in Indian children in Britain or elsewhere. Certainly none were identified for children of any ethnicity from the open-ended narratives in B-CAMHS (Robert Goodman – personal communication).

Qualitative and anthropological evidence

Assessing non-Western populations using Western-derived symptoms or classificatory systems may cause one to miss locally important aspects of mental health experiences. It is here that emic studies can be most informative. Such studies develop culturally meaningful typologies from scratch, and only later compare these to etic constructs [182, 206]. While most emic studies are on adult mental health [e.g. 207], a few focus on children [208-209].

I know of no emic assessments of Indian child mental health, either in Britain or elsewhere. There have, however, been a number of qualitative studies of adult mental health concepts among different ethnic groups in the UK. These do not indicate marked differences between ethnic groups. For example, O’Connor *et al.* [210] conducted in-depth interviews with 116 men and women of Bangladeshi, Black-Caribbean, Indian, Pakistani White British and White Irish ethnicity. Although not identical, there were substantial similarities across groups in descriptions of the experience of emotional distress. Likewise two emic studies of ‘thinking too much with my heart’ [211] and ‘sinking heart’ [212] in British

Punjabi adults concluded that these phenomena have much in common with DSM-IV and ICD-10 definitions of depression. There were, however, also some important differences. For example, some core depressive symptoms (e.g. loss of self-worth) were absent, while heart-related symptoms were seen as central and received considerable attention.

There is less qualitative research regarding child mental health, but what does exist likewise provides no evidence of major systematic differences. One study of Scottish 12-14 year olds reports that there were no striking differences between the 120 White children and the 25 minority ethnic children (mainly Pakistanis) in how they thought of 'mental health' [213]. Likewise a study of 60 parents of White, Black, South Asian and Mixed race children in London found broad similarities across groups in the feelings and behaviours identified as problematic in children [214].

Relation of psychiatric classes to external predictors

There is therefore some evidence of the validity and cross-cultural validity of the emotional/behavioural/hyperactivity dimensions of child mental health problems. Further support lies in the relation of these domains to external correlates. As reviewed in Section 2.2.4 Chapter 2, domain-specific associations exist for many risk factors. The same applies to many other factors including aetiology, heritability, prognosis and treatment response [reviewed in 167]. For example, compared with behavioural disorders, hyperactivity shows earlier onset, higher heritability and a greater likelihood of successful pharmacological treatment.

Unfortunately, few studies from low and middle income countries include data on child, family and area factors. Gender is one of the few routinely reported variables. A recent review reports that in almost all populations ever studied around the world, boys are more vulnerable to externalising problems and girls to internalising problems [191]. Those studies and reviews which do consider other psychosocial factors likewise suggest important similarities in mental health correlates [e.g. 42, 77-78, 215]. Some commonality of molecular, cognitive and perhaps even social mechanisms is also indicated by the apparent efficacy in low income countries of Western medications and cognitive behavioural therapies, although systematic evidence on this point is lacking [205, 216].

Conclusion

In this PhD I focus on broad domains of common child mental problems and do not use individual diagnoses such as 'generalised anxiety disorder' as outcomes. This is worth emphasising because the most convincing critiques of cross-cultural comparisons relate either to rare culture-bound syndromes or to the validity of fine distinctions (e.g. depression vs. anxiety) within larger categories [14, 182]. The evidence presented above suggests these critiques may be less relevant to the broader domains of emotional, behavioural and hyperactivity problems.

This does not deny the importance of culture. On the contrary, some of the most convincing demonstrations of cross-cultural similarities are those which also identify important differences. For example, Luk *et al.* [203] present considerable evidence for a 'culture free' core of hyperactivity behaviours in all cultures but simultaneously document various 'non-culture free' elements. These elements are imposed on top of the culture free core, and affect the development, recognition and management of hyperactivity.

Yet I also think it important not to overstate the influence of culture. One criticism of much cross-cultural psychiatry is a tendency to ignore the interdependence and overlap between cultures, an interconnection which is increasingly prominent after several decades of globalisation [217-218]. This applies particularly strongly to studies which, like this thesis, compare ethnic groups from the same country.

A research interest into cultural differences should therefore not blind one to the potential for similarity. Although less extensive than would be ideal, the evidence presented above provides no *a priori* reason to believe that the common child mental health problems will be inherently incomparable across British ethnic groups. In Section 8.2 Chapter 8 I build upon this conclusion by assessing this issue empirically in B-CAMHS.

2.4 Measurement of mental health problems

In discussing the possibility of fundamental differences in child mental health problems, Section 2.3 focussed on whether cross-cultural comparisons had the *potential* to be meaningful. Yet even if a comparison is potentially meaningful in principle, this does not mean that it will be fair in practice. Rather systematic differences in how mental health problems are recognised or reported may still lead to biased conclusions. As in the previous section, these measurement issues reflect challenges which apply in all child psychiatry and which may jeopardise comparisons between populations of any sort. I outline these challenges below, and review some techniques which can improve the validity of mental health assessment.

2.4.1 The centrality of the subjective experience

One implication of the phenomenological approach to classification is the central role played by the subjective feelings and perceptions of children and their carers. This is firstly because while some symptoms can be measured objectively (e.g. physical restlessness), many others refer to internal psychological states (e.g. anxiety) or interpersonal interactions (e.g. defiance). Furthermore, all symptom-clusters only constitute a problem or disorder if they cause substantial distress and/or impairment. The subjective experience is therefore integral to the definition of mental health problems and disorders, and any meaningful approach to measurement must recognise this [219].

This is not unique to the study of mental health. For example, Midgeley [220] makes a parallel argument regarding the scientific study of consciousness when she argues that no account of consciousness can be adequate if it does not retain consideration of the subjective experience. Likewise, although physiological tests form gold standard measures in other areas of medicine, I believe that purely objective assessments could never provide a fully adequate way of capturing mental health. Instead meaningful assessment necessarily requires a detailed account of symptoms (many of which are subjective) and some understanding of what those symptoms mean for the child.

This poses substantial challenges for achieving comparability of mental health measurement across individuals. Such comparability is important in epidemiological surveys, and avoiding systematic differences is particularly crucial for comparisons between groups. Assessing symptoms and impact in an accurate and comparable way is therefore central to the validity of mental health measures. The absence of objective gold standards means one can never conclusively demonstrate the superiority of a given method. Instead, just as for classification systems, one needs to build a case for the greater validity of some approaches compared to others. Likewise, one needs to build a case for the validity of specific mental health measures despite the absence of a gold standard. Below I review some of the principal techniques for achieving this, after first setting the discussion in context by outlining some key sources of bias which may undermine comparability between individuals.

2.4.2 Mechanisms of bias in mental health assessments

<u>Selection bias</u>

Selection bias occurs when participants in a survey are systematically different from non-participants. This is a particular cause for concern in this PhD given that many British surveys report lower participation rates for minority ethnic individuals or in minority ethnic areas [221-222]. I return to this issue in Section 6.3 Chapter 6.

<u>Information bias</u>

Different thresholds for endorsing items

A deviation from criterion equivalence occurs if children, parents, or teachers differ in their threshold for reporting symptoms or impact [203, 223]. This can either occur for specific items (an item bias) or across most or all items (a method bias) [224].

Individuals inevitably show some differences in their thresholds for endorsing subjective, self-reported outcomes. That these differences may be substantial for mental health measures is indicated by the large informant-specific effects which most measures show [225]. For example, Achenbach *et al.* [226] report a meta-analysis of studies in which different informants completed questionnaire measures about the same child. Even when the informants had a similar relationship to the child (e.g. mothers and fathers, or pairs of teachers), Pearson's correlation coefficients were only around 0.6. This is substantially lower than the test-retest reliabilities of 0.7-0.9 typically reported when the informant is kept constant (e.g. the same teacher two times) [227].

Large informant-specific effects are problematic because they indicate that much of what a measure is capturing is the viewpoint of a particular informant. Some of these differences in informant viewpoint may in fact have substantive causes – for example, a child may be oppositional with one parent but not the other. This would be an example of *person*-specificity in the manifestation of mental health problems, analogous to the more commonly considered situation-specificity (e.g. problems at home but not at school). Nevertheless, much of the informant effect may have no relation to the child's symptoms but rather simply reflect that particular informant's attitudes or expectations. If so, then at the very least this will introduce measurement error. Moreover, if individuals from different groups differ *systematically* then this may mean that disorders are missed or misidentified more frequently in one group than another. Likewise, given that most dimensional measures are created by summing a finite number of symptoms, systematic threshold differences could create spurious differences in mean scores.

This possibility represents a fundamental threat to the validity of any inter-group comparison. For example, there is some evidence that parental reports of a child's mental health differ systematically according to the parent's own mental health [228] or that of their partner [229]. If so then this could lead to misleading conclusions regarding the degree of association between the mental health problems of parents and children. Such biases are likewise a major concern in cross-cultural comparisons, as highlighted in several reviews [201, 205]. One possible example of such as bias is the observation that Hong Kong teachers seem to have a lower threshold for identifying hyperactivity problems than British teachers [203]. Another possible example comes from a study which my colleagues and I conducted comparing parents and teachers from Norway and Britain. Our findings indicated that, compared to their British counterparts, Norwegian parents and teachers had a higher threshold for identifying emotional difficulties in children ([230]; see Appendix 3). Variation has even been observed among mental health professionals using uniform assessment criteria, as indicated by comparisons of hyperactivity ratings in the USA, Indonesia, China and Japan [231].

Less work has been done comparing ethnic groups within Britain. One recent exception is an investigation from East London of teachers' reporting of hyperactivity in South Asian and White children. As judged by objective movement measurements, teachers seemed to overestimate the hyperactivity of South Asians; the South Asian group rated 'overactive' had movement scores no greater than their 'normal' White peers [232]. Other studies do not directly investigate potential biases of this sort, but do present findings which suggest they are plausible. For example, Hackett [10] reports that behaviours which White parents saw as 'healthy self-assertion' (e.g. snatching back a toy grabbed by another child) were often seen by Gujarati Indians as worryingly aggressive. This might lead to a lower threshold in the Gujarati parents for reporting behavioural problems, a bias which would tend to over-estimate problems in Indians compared to Whites.

Different amounts of disclosure

The above studies consider disclosure to pre-specified closed questions. In clinics and in some epidemiological studies, responses to open-ended questions are also important in making diagnoses. Inter-group differences in the length or content of these responses are therefore another potential source of systematic bias. Plausible causes for such differences include variation in the respondent's fluency in the language of interview, the respondent's trust in the interviewer, or the interviewer's use of prompts and active questioning.

2.4.3 Approaches to identifying and addressing information bias

<u>Structured diagnostic interviews vs. questionnaires</u>

One method of increasing inter-informant comparability is to ask questions which are more detailed and more specific, as these reduce the scope for individual interpretation. This can be achieved by administering structured diagnostic interviews rather than brief questionnaires; while both are fully reliant upon responses to closed questions, the questions in the latter are typically much more numerous and more detailed.

That structured diagnostic interviews may enhance cross-cultural comparability was supported in our above-cited comparison of how British and Norwegian adults rate child mental health ([230]; Appendix 3). The apparent under-reporting of emotional problems by Norwegian teachers and parents on brief questionnaires was not observed when the same informants completed detailed structured diagnostic interviews. This implies that detailed interviews may be less prone to bias than brief questionnaires, and that comparing the two is one method for identifying information biases in questionnaires.

Fully-structured, unstructured and semi-structured interviews

A further important strategy is to include open-ended questions in which respondents describe or elaborate on areas of concern. The lack of open-ended sections is an important source of measurement error in fully-structured respondent-based clinical interviews. Such measurement error arises because closed questions are prone to misunderstandings around symptoms with colloquial uses (e.g. obsessions), and fully-structured interviews consequently tend to over-diagnose conditions like obsessive-compulsive disorder [233]. Such measures may simultaneously miss the 'bigger picture' and fail to identify children who merit a Not Otherwise Specified diagnosis [34, 172]. Fully-structured measures may also overlook or misidentify less common disorders such as anorexia, autism, or psychosis which can be difficult to capture using closed questions but which have such distinctive symptoms that verbatim descriptions are often unmistakable [234].

It is for these reasons that investigator-based interviews are generally regarded as more valid than respondent-based interviews (for a review of both forms of interview, see [227]). Investigator-based interviews provide a list of symptoms to be covered, but the investigator then uses flexible questioning until sufficient information about each symptom has been collected to make a diagnosis. This provides substantially greater scope for clarifying misunderstandings and for ensuring the identification of all problems which are of subjective importance to the child. This, in turn, increases the likelihood that diagnosed disorders correspond to genuine mental health problems, and vice versa. A further advantage is that if the same investigator administers multiple measures then this may increase the comparability of diagnoses across children. These advantages may all be particularly critical when making comparisons across cultural groups [206].

The disadvantage is that investigator-based methods require clinicians or other highly trained individuals, and may therefore be prohibitively expensive in large studies. One novel approach used in B-CAMHS was to supplement a fully-structured interview with open-ended questions. Informants were prompted to expand on any areas of concern in these questions, and their responses were later reviewed by a child psychiatrist. As argued in more detail in Section 5.3.3 Chapter 5, this semi-structured approach may have incorporated many of the advantages of investigator-based methods into a format that was feasible for a large epidemiological study.

Multiple informants in making diagnoses

A further common method for increasing the validity of child mental health diagnoses is to incorporate information from multiple sources. Frequently used sources include direct observation in clinical or naturalistic settings, reports by adults (particularly parents and teachers) and questioning of the child [227].

Multiple informants are important because parents, teachers and children vary systematically in the quality of information they provide about different problems. Teachers and, to a lesser extent, parents are usually far better at identifying externalising symptoms such as defiance or restlessness than internalising symptoms such as misery [2-3, 34, 235]. Conversely, children often show little insight into whether their behaviour is overactive, disruptive or inappropriate [236]. This lack of insight may in part be a symptom of the pathology. It also, however, seems likely to reflect the general difficulty that children (especially younger children) have in making accurate, objective judgements about themselves compared with others [e.g. 237]. Most diagnostic interviews therefore only interview older children and, for externalising symptoms, only ask about very specific behaviours such as theft or fire setting.

Collecting information from multiple informants therefore has the benefit of permitting principled decisions about whose information to use for a particular sort of problem. Moreover, *triangulating* between informants may provide an even more complete picture than is possible from any informant individually. First, details from one interview may clarify or extend information from another. Second, children inhabit multiple non-overlapping worlds and different informants will offer insights into problems in each setting. For example, most parents know more about a child's feeling and behaviour at home than at school, and vice versa for teachers. As for the internal dynamics of peer groups, only the child may be able to give an informed perspective. High quality information across settings is particularly crucial for disorders like hyperactivity where the diagnostic criteria require symptoms to be pervasive.

Considering data from across multiple informants can also be valuable in cross-cultural comparisons. This is illustrated by a study in inner London in the 1970s in which teachers rated Black Caribbean children as more disruptive than Whites [238]. Teachers did not report worse peer relations in the Black Caribbean children, however, despite the fact that poorer peer relations typically accompany behavioural disorders. Black Caribbean children were also not rated as more disruptive than average by their parents. Without further investigation, it is not possible to determine whether these results reflect a reporting bias (e.g. teachers had a lower threshold for reporting behavioural problems in Black Caribbean children) or a substantive factor such as situation-specificity (e.g. the Black Caribbean children were only disruptive at school). Regardless of the cause, the authors conclude that in the Black Caribbean sample, disruptive behaviour at school did not necessarily have the same clinical implications as a core symptom of behavioural disorders. This highlights the importance of triangulating evidence from across multiple informants and across multiple domains of functioning. In Chapter 7 I address this in relation to the comparison of Indians and Whites.

Challenges in interpreting disagreement between informants

The purpose of collecting data from parents, teachers and children is that they may differ in their knowledge, perspectives or insight. Yet while collecting information from multiple sources therefore has the potential to enhance validity, it is also in tension with the common desire to make a single global assessment. As with other challenges in this section, this issue applies not only to decisions about individuals (e.g. does this child merit a diagnosis of depression?) but also to comparisons between groups. For example, this PhD is motivated by the apparent Indian advantage for multi-informant clinical diagnoses. If this advantage were reported by parents, teachers and children alike then this would certainly strengthen the conclusion that Indian children really do enjoy a mental health advantage. Yet if the advantage were not consistently observed across informants it would *not* necessarily follow that the advantage was not real – instead, it might be that the advantage was real but was confined to a particular setting. The possibility of multiple interpretations of a discrepancy between informants is precisely why it is unclear how to interpret the greater teacher-reported disruptive behaviour in Black Caribbean children in the study described above.

Collecting information from multiple sources therefore necessarily carries the challenge of dealing with disagreement. This challenge is particularly acute because in fact informants typically show low agreement in child mental health assessments [227]. For example a review and meta-analysis gave Pearson correlation coefficients of 0.27 between parents and teachers (41 samples), 0.25 between parents and children (14 samples) and 0.20 between teachers and children (21 samples) [226]. Little is known about the relative contribution of alternative possible explanations such as measurement error, reporting bias or genuine situation-specificity [239].

The most suitable method for dealing with imperfect agreement may depend on the nature of the discrepancy and the purpose of the evaluation. Clinicians often use simple algorithms such as assuming a symptom is present if any respondent mentions it, and there is some evidence that that this works as well as more complicated methods [240]. Ideally these simple algorithms will be supplemented by more sophisticated judgements about how much weight to give different informants based on the attitude of the respondent, the context of the interview or the kind of problem they are reporting about (e.g. giving more weight to a child's report of anxiety symptoms than hyperactivity symptoms).

Epidemiological studies can also use such techniques if clinical judgement plays some role in assigning diagnoses. Epidemiological studies may also combine information across informants using algorithms or statistical models (e.g. latent class modelling). Other possible techniques include conducting psychometric evaluations of the data from different informants in order to make informed judgements about the validity of the data from each source; and presenting sensitivity analyses using data from different informants. These latter two strategies are the ones which I use in this PhD.

2.4.4 Cross-cultural comparisons: sources of bias

The differential participation, threshold biases and differential disclosure discussed in Section 2.4.2 are general *mechanisms* whereby selection and information biases may jeopardise comparisons. For specific comparisons, it may also be informative to study plausible *sources* of these biases. As described below, language and type of parent informant are two factors which might be the source of systematic differences in the responses of different ethnic groups in the UK. I therefore assess these factors directly in Section 8.3 Chapter 8.

Language of interview

Variation in how a question is understood is one cause of deviation from criterion equivalence. Understanding questions in unintended ways is frequent in all psychiatric epidemiology [233], but is particularly likely when using translated instruments or when the respondent is not fluent in the language of interview. A lack of fluency may also cause individuals to provide less information to open-ended questions.

In Britain, many minority ethnic adults do not speak English as a first language [241]. For example, while over 99% of White and Black Caribbean families in the Millennium Cohort Survey spoke English at home, this was true of only 89% Indians, 88% Black Africans, 76% Pakistanis and 67% Bangladeshis [242]. A substantial minority of parents from these groups may therefore have found language a real barrier to participating in B-CAMHS, which could lead to selection bias. Alternatively, misunderstandings or poor translation may have lead to information bias between respondents who were fluent in English and those who were not. This is one explanation for the finding in two large adult surveys that a South Asian mental health advantage was confined to participants with poor English [243] or interviewed in languages other than English [244].

Type of informant

Poor English might also lead mothers to decline to participate and fathers to respond instead. If so then language would be indirectly implicated in a second potential source of bias. An over- or under-representation of fathers might also occur for other reasons. These include cultural factors such as the perceived appropriateness of a woman being interviewed in private; socio-demographic factors such as the prevalence of single-parent families; or socio-economic factors such the prevalence of male economic inactivity. Cross-cultural bias could then result if fathers differed systematically from mothers in their knowledge or perceptions of their child's emotions and behaviours.

2.4.5 Validating particular mental health measures

I have so far focussed on general strategies for increasing validity and/or investigating bias. Ultimately, however, all mental health assessments rely on particular questionnaire and interview measures. In validating these measures the absence of a gold standard is again a challenge, as it prevents tests of criterion validity.[6] One therefore instead needs to demonstrate the *construct* validity of a measure relative to the existing framework of theoretical and empirical evidence [227, 245]. As when validating classification systems, this typically involves presenting both internal and external evidence of validity. These include:

- Examining whether the measure's symptom structure conforms to theoretical constructs (e.g. internalising and externalising problems).
- Comparisons with established measures to evaluate convergent validity (high correlation with measures of the same construct) and discriminant validity (little or no correlation with measures of different constructs).
- Verifying that higher symptom scores are associated with greater impact.
- Group differentiation, e.g. how well a measure distinguishes clinic and community samples of children.

[6] Criterion validity refers to validation relative to a perfect measure (a gold standard), while construct validity refers to validation relative to other imperfect measures. See Table 13.2, Appendix 1 for a general overview of these and other forms of validity.

- Hypothesis testing, e.g. whether a measure produces prevalence rates or risk factor associations similar to other studies, or whether the measure shows improvement after a treatment of known efficacy

There is clearly some risk of circularity in 'bootstrapping' the validity of new (imperfect) measures based on existing (imperfect) measures, theory and empirical findings. For example, a new measure may show high correlation with existing measures or have a risk factor profile consistent with the existing literature, but this is only evidence of validity if the existing measures and literature are themselves valid. Nevertheless, the absence of a gold standard means that marshalling multiple lines of evidence to indicate construct validity is the best that is usually possible. This underpins my approach when summarising existing evidence on the B-CAMHS measures (Section 5.3 Chapter 5) and when presenting further original analyses of their validity in Indians (Section 8.2 and 8.3, Chapter 8).

2.5 Summary and conclusions

To summarise, the common child mental health problems and disorders share two features which create particular challenges for definition and assessment. The first is a phenomenological approach to classification which is based upon the observation of symptom-clusters and impact, and not upon an understanding of underlying disease processes. This prevents definitive demonstration of the validity of a proposed classification system, and means that systems which are valid in one population may not be valid in another. One therefore has to build a case for the validity of a classification system, and to do so in every population to which it is applied. The second feature is the centrality of subjectivity to both mental health symptoms and impact, and the resulting absence of fully objective measures. This again means one must build a case for the validity of particular measures in particular populations.

These challenges of validating classification systems and measures cut across all psychiatric epidemiology, but become particularly acute in comparisons across time, space or culture. These issues are therefore particularly relevant to the ethnic comparison motivating this thesis. The evidence reviewed above suggests that in general there are substantial cross-cultural similarities in child emotional, behavioural and hyperactivity problems. Nevertheless, the absence of research specifically on Indian children highlights the importance of examining this issue directly. Moreover, there is some evidence that parents and teachers of children from different ethnic groups may differ in how they report symptoms, or may present discrepant accounts. This highlights the importance of documenting precisely what ethnic differences are reported by each informant; of investigating both the mechanisms and the sources of potential biases; and of thereby trying to understand the causes of any inconsistencies. These considerations are central to the first two aims of my PhD, and motivate the analyses presented in Chapter 7 and Chapter 8.

Chapter 3 Ethnicity in Britain: conceptual issues and empirical evidence

Having discussed child mental health as an outcome, I now turn to ethnicity as an explanatory variable. First I present general principles relevant to any epidemiological study, although with a focus upon their application in Britain. I then review the migration histories and current characteristics of the main minority ethnic groups in Britain. Together this provides the context for the systematic review in the next Chapter of ethnicity and child mental health in the UK.

3.1 Ethnicity as an explanatory variable

The past three decades have seen ethnicity and health become a major topic of epidemiological investigation. Several commentators have, however, criticised much of the existing literature for inadequate theorisation and unsophisticated analyses [246-248]. The central problem is a tendency to treat ethnic groups as natural categories which are fixed in themselves, uncontested in their membership and therefore straightforward to analyse.

In fact there is no consensus about how to conceptualise ethnicity or about whether it is a useful focus of epidemiological study. Measuring ethnicity in a way which is reliable and valid also poses important challenges, as does analysing ethnic differences in a way which is maximally informative and which adequately controls for confounders such as socio-economic position. I address these points below, drawing wherever possible on examples from mental health.

3.1.1 Conceptualising ethnicity

<u>What is ethnicity?</u>

Nazroo [249] conceptualises ethnicity as comprising two broad axes; ethnicity as identity and ethnicity as structure. Drawing additionally upon theoretical models developed in gender studies [250] I have extended this to create the following framework:

- **Biological bases for ethnic differences:** Scientific attempts to divide humans into biologically discrete races were prominent in Europe and American from the eighteenth century [251]. This endeavour was widely rejected following the Second World War [252], and population genetics has since falsified the existence of human races in the biological sense of subspecies [e.g. 253]. There exists, however, a very active research programme into the possible health implications of ethnic variation in allele frequencies or epigenetic effects [254-256]. Such effects undoubtedly contribute to some ethnic differences in health, although for most outcomes the magnitude of the contribution is highly uncertain.

- **Ethnic identity:** An individual's sense of their own ethnic self is often rooted in the feeling of belonging to a personally meaningful collectivity. The strength of this sense of membership may be shaped by a group's current context and historical experiences. It is also internally generated through distinctive cultural traditions, shared modes of thinking and behaving, and actual or symbolic links with a place of origin [247, 257-258]. The importance of these dimensions in capturing how British individuals see their ethnic identity is indicated both by reviews of the qualitative literature [259] and by factor analysis of a recent nationally representative survey [260].

- **The social and structural basis of ethnicity:** Finally, ethnicity can be understood as a social phenomenon generated and perpetuated by a group's structural position in society [249]. This includes factors such as a group's physical/geographical location, their socio-economic position, and their degree of political representation. It also includes the group's socio-cultural significance, including the attitudes and behaviours towards that group at a societal level. For many minority groups in Britain, the most salient aspect of this is the widespread experience of interpersonal and institutional racism [241, 261]. Racism may also intersect with other aspects of the social basis of ethnicity, for example by contributing to geographical segregation. Anti-racist perspectives frequently emphasise the social basis of ethnicity, highlighting the racism, power inequalities and socio-economic disadvantage which characterise the ‘structure’ facing many minorities [246, 262-263].

These three dimensions of ethnicity are not alternatives. Rather they represent different aspects of a complex phenomenon, and each dimension may be more or less salient in different contexts or for understanding different aspects of ethnicity. In investigating the apparent mental health advantage of Indian children, I focus throughout this thesis upon ethnic identity and the social basis of ethnicity. I nevertheless included biological factors in my typology because, as outlined below, I believe they could in theory be relevant and that their absence therefore deserves consideration and justification.

Are biological aspects of ethnicity relevant for this thesis?

That variations in allele frequency underlie ethnic differences in some rare health outcomes is not controversial – well-known examples include sickle-cell anaemia in Black Africans or cystic fibrosis in White Europeans. As summarised previously in Section 2.2.5 Chapter 2 it is also clear that genes affect the mental health of individuals. Many mental disorders or mental health-relevant traits show high heritability [159], and convincing examples have recently been put forward for moderate effects of common alleles [165].

Nevertheless, I believe that there are several reasons for believing that genetic differences are not the most plausible explanation of the Indian advantage. On the one hand is the evidence from studies modelling the fall-off of risks to relatives or conducting meta-analyses of genome scans. These indicate an oligogenic and perhaps polygenic genetic basis for all the common mental disorders and mental health traits examined thus far [264]. That is, most observed heritability seems likely to stem from relatively small contributions (odds ratios usually <2) from relatively large numbers of genes (at least 5-20 loci). Simultaneously, within-population genetic variation is far greater than between-population variation [253], and what between-population variation does exist is characterised by gradients (‘clines’) between regions rather than discrete clusters (‘clades’) [265]. The discovery that human populations are more genetically structured at the micro, meso- and macro-levels than initial estimates indicated [266-267] does not change this basic conclusion that continental and sub-continental populations are not genetically bounded.

This does not make it impossible that genetic factors contribute to the Indian mental health advantage. After all, obesity and type 2 diabetes likewise have a complex genetic basis, yet genetic differences seem likely explain at least part of the excess risk in South Asian adults [268-270]. Nevertheless, the *combination* of multiple genetic influences and an absence of discrete 'breeding populations' makes it less likely that a genes favouring good mental health would arise in Indians through random processes such as founder effects or genetic drift. The non-random process of natural selection is a more plausible mechanism from this perspective, but I know of no reason to hypothesise stronger selective pressures for child mental health genes in India than in other populations. As for non-genetic biological effects, evidence for mental health outcomes is limited to rare disorders like schizophrenia and extreme environmental exposures like maternal famine [271-273].

In summary, I know of no research which addresses the question of whether there could be a genetic basis for the apparent Indian child mental health advantage. There are, however, several reasons to think that this is comparatively less plausible than social or cultural explanations. This conclusion, and my reasons for it, parallel Krieger's more general critique of the assumption of that most observed ethnic differences in health have biological causes [274]. As such, I believe that research into biological causes for the Indian mental health advantage would be far better justified if it were demonstrated that the difference could not readily be explained by known environmental factors. Investigating whether this is the case is, of course, the purpose of this PhD.

Researching ethnic variation – minimising the potential for harm

The potential for harm

A worrying potential continuity between current health research into ethnicity and past scientific racism is the danger of doing harm. Concentrating on ethnic *differences* may obscure similarities between groups, and result in sub-optimal and divisive uses of resources (e.g. ethnically-targeted rather than mainstream services). Focussing on rare 'ethnic' problems may also distract from public health priorities among minority ethnic groups [247] – particularly if an ethnocentric perspective leads to any problem which is no more common than in Whites being dubbed unimportant. An ethnic focus also risks essentialising health differences as being located genetically, physiologically or culturally within the minority ethnic group [275]. This in turn may cause adverse life circumstances and structural disadvantage to be mistaken for cultural difference [276-278].

These possible harms are compounded by a research focus upon problems in minority ethnic populations. For example, far more studies have investigated the high rates of schizophrenia in Black Caribbeans than the low rates of schizophrenia in South Asians [14, 279]. Furthermore, while excess problems are often interpreted as reflecting pathological cultural characteristics, the converse is not true. For both physical diseases such as rickets [246, 280] and mental disorders like schizophrenia [14, 243, 281] the cause of excess rates is frequently located within a particular culture, often in a manner closely linked to popular stereotypes. By contrast, lower rates of mental health problems are commonly interpreted as possible artefacts (e.g. healthy migrant effects, underutilisation of healthcare, atypical symptom presentation) or else allow only a palliative role for culture (e.g. social support systems replace specialist services when problems develop). For an example of all of these interpretations, see Cochrane and Bal [282]. Moreover, any protective role granted to culture often again draws on popular stereotypes [283]. Thus South Asian mental health is often explained without evidence in terms of characteristics of the 'Asian family' – low rates being attributed to the extended family's secure environment, high rates to the same extended family's repressive attitudes [284].

Maximising research benefits

None of this means that epidemiologists should ignore ethnicity. To do so would be to disregard powerful social categories which may have important and distinct consequences for psychological, inter-personal and material well-being [246, 248, 285-287]. The solution is therefore not to abandon ethnicity research, but rather to ensure it is conducted to a high standard. First, in the absence of consensus, one's theoretical conceptualisation of ethnicity must be stated explicitly. Secondly, clarity of research purpose is necessary to achieve a thoughtful empiricism which is not simply driven simply by data availability [246, 287-292]. This is of course vital in all epidemiology, but is particularly crucial given the routine collection of ethnicity by the UK censuses, NHS and other public bodies. Finally, as discussed subsequently in Section 3.1.3, care must be taken not to undermine theoretical sophistication with crude methods of measurement or analysis.

3.1.2 Measuring ethnicity: the UK census classification system

Most large-scale surveys ignore the complex and context-dependent nature of ethnicity, instead operationalising it as a relatively small number of pre-specified, discrete groups. This is often done in an *ad hoc* manner which uses terminology inconsistently, conflates race and ethnicity,[7] and even invents ethnic group (e.g. 'Urdus', defined by language) [287-292]. A review of the American Journal of Epidemiology and the American Journal of Public Health found that while 77% of recent articles mentioned ethnicity, there was an enormous diversity in the terms used and the number of groups recognised (from zero to 24) [293]. This inevitably creates problems of non-comparability across datasets. Articles also frequently failed to describe how ethnicity was assigned.

In Britain, the introduction in the 1991 census of an ethnicity question (see Box 3.1) has had a major stabilising effect. Self-assigned ethnicity using the census response options is now used in most research studies [294], including B-CAMHS (see Section 6.2.3 Chapter 6).

Box 3.1: Ethnic response options in England and Wales in the 1991 UK census

- White
- Black Caribbean
- Black African
- Black – Other
- Indian
- Pakistani
- Bangladeshi
- Other – Asian
- Chinese
- Other

Source: Simpson et. al. 2006 [221]. Response options differed slightly in Scotland.

This standardisation clearly has major potential benefits for research, as does the increasingly widespread collection of ethnicity data. Yet the availability of large, comparable datasets makes it all the more necessary to retain a critical perspective regarding this particular method of operationalising ethnicity [287]. Conceptually, the

[7] Originally, biologists used race to describe subspecies, and anthropologists used ethnicity to describe cultural groups. This distinction is no longer apparent, however, and increasingly race means in the US what ethnicity means in Britain. The terms are also often used interchangeably or else replaced with ill-defined compounds (e.g. race/ethnicity) and sanitised alternatives (e.g. 'race').

census combines a haphazard mix of principles, including skin colour (e.g. ‘White’), regional origin (e.g. ‘Black African’) and nationality (e.g. ‘Pakistani’). In its favour, this does appear to reflect meaningful collectivities for people in Britain [241]. Indeed, these response options were to a large extent driven by consumer testing so as to maximise their acceptability and minimise confusion [288]. Yet precisely for this reason, the census categories are not neutral. Rather they embody a long-standing and distinctively British socio-political model in which only minority groups are ‘ethnic’ (hence the originally undifferentiated White group[8]) and in which ethnicity is manifested in an unstable mix of skin colour and culture [15].

The census categories therefore embody a particular model of ethnicity. Moreover, their routine administration also perpetuates this model by fostering the idea that this is the best – or indeed the only – approach to classifying ethnicity. Subsequent analyses are then in turn constrained to use these categories, thereby legitimating them further.

Yet the census categories may not always ‘carve up’ ethnicity in the most informative ways. For example, ethno-religious differences within South Asians (e.g. Hindu, Muslim and Sikh) may be more important than ethno-national differences for personal ethnic identity [262] and for various aspects of family and socio-economic life [295]. The census categories may also often hide considerable variation. This is illustrated by variation in the education achievements among White minority children from different regions of Europe and among Pakistani children from different parts of Pakistan [296]. The considerable heterogeneity among Indian migrants (see Section 3.2) makes internal diversity particularly plausible in this group. I cannot address these questions in B-CAMHS, however, which like so many other surveys only measured ethnicity using the census categories. This absence of information on alternative or complementary aspects of ethnicity is therefore an important limitation in this thesis.

[8] In the 2001 census the ‘White’ group was expanded somewhat to include ‘White British’, ‘White Irish’ and ‘White Other’.

3.1.3 Ethnicity as an analytic variable

Ethnic variation is a starting point not an explanation

Some authors argue that ethnicity cannot meaningfully be treated as a cause because it is an attribute which does not allow for counterfactual states (e.g. a Black person not exposed to 'Blackness') [297-298]. I do not agree with this position, nor do I believe ethnicity is a single, non-modifiable attribute. For example, while a Black person may necessarily be exposed to 'Blackness', this Blackness need not include (for example) the experience of discrimination in the labour market.

Nevertheless, I certainly agree with Rutter [299] that multifaceted, non-homogenous concepts like ethnicity only become meaningful explanatory variables when broken down into their constituent parts [see also 202]. Moreover, different factors may be relevant for different ethnic groups or for different health outcomes. Researchers must therefore not confuse the *description* of ethnic variation in health with an *explanation* for that variation [287, 289, 292, 300-301]. Instead, the observation of any ethnic difference should be a starting point for further investigation into operative causal mechanisms. Identifying these mechanisms may then enable the imagining of counterfactual states, and so inform the design of public health interventions.

Of course, this situation is not unique to ethnicity but applies to all complex and multifaceted socio-cultural and economic factors. For example, variation between manual and non-manual occupational groups or between urban and rural residents is not informative in itself; there will rarely be anything intrinsically 'manual' or intrinsically 'rural' about health differences. Rather further research into causal mechanisms is necessary to understand why such differences exist.

Causes of ethnic variation should be directly measured not assumed

Causes of ethnic variation in health may include differences in allele frequency; lifestyle and cultural factors; socio-economic position (SEP); or how individuals and groups are treated by society. It is therefore crucial to investigate causal mechanisms directly, by testing key assumptions and by comparing and contrasting alternative explanation. Again, this is not unique to ethnicity but applies to all multi-faceted phenomena such as manual occupation or rural residency.

Yet despite the multiple dimensions of ethnicity, several commentators have lamented a tendency to assume without evidence that ethnic differences have genetic or cultural causes [246-247, 302]. This risks pathologising minority ethnic groups and may ignore other structural sources of disadvantage such as racism or low SEP [290, 303]. The importance of considering multiple axes of disadvantage is illustrated by the 1999 Health Survey for England, which showed greater health disparities between different income tertiles in the same ethnic group than between different ethnic groups in the same tertile [304]. Section 3.1.3 therefore focuses on the necessity and the challenges of considering socio-economic inequalities when analysing ethnic differences.

Analysing socio-economic position and ethnicity

In Britain, SEP is strongly associated both health [305] and ethnicity (see Section 3.2). Yet while SEP is therefore an important potential confounder when comparing ethnic groups, controlling for SEP is far from straightforward. Of course, this is true to some extent across epidemiology. It is particularly acute here, however, because migration and membership of a minority ethnic group may result in different facets of SEP being 'pulled apart'.

For example, downward social mobility upon migration may result in minority ethnic individuals having above-average educational qualifications at any given income level. Given the protective effect of parental education against child mental health problems, adjustment for income alone might result in substantial residual confounding. In this case the nature of the residual confounding would be to create a misleadingly favourable impression of the mental health of minority ethnic children relative to Whites. Conversely, controlling only for parent education might under-adjust for material deprivation and create a misleadingly unfavourable impression of minority ethnic child mental health. Under-adjustment for SEP may also occur through ignoring the cumulative effects of multiple disadvantages, if these are disproportionately common among minority ethnicities. I present emprical evidence on these points in Section 3.2.2.

Using single, crude measures of SEP when comparing ethnic groups may therefore lead to substantial residual confounding, and this may either exaggerate or mask ethnic differences. Multiple or composite indicators are one way to improve adjustment for SEP in any epidemiological study [306], and may be useful when studying minority ethnic groups [307]. Yet even with such indicators, there remain formidable challenges in measuring SEP accurately and dealing adequately with the complexity of the relationship between SEP and ethnicity. This warns against accepting too readily that one has ever fully adjusted for SEP when comparing ethnic groups.

Finally, even if SEP proves central to explaining ethnic differences in health, this does not mean that ethnicity is irrelevant or nothing remains to be 'explained'. For one thing, the socio-economic disadvantage of many minority ethnic groups may partly reflect aspects of their minority ethnic status. These might include language barriers to employment, exclusionary racist practices, or a preference among minority ethnic individuals for investing in assets in their country of origin. Moreover, socio-economic inequalities are themselves like ethnic inequalities in being a starting point for further analysis rather than an explanation in their own right.

What is therefore needed is an approach which examines in detail the association between ethnicity and multiple indices of SEP, and which investigates the individual- and family-level mechanisms underlying any variations in health. One major aim of this thesis is to use B-CAMHS to develop precisely such an approach for child mental health in Britain.

3.2 Ethnic groups in Britain

Any adequate account of ethnic differences should therefore examine underlying causal mechanisms directly. This requires one to identify and measure the multiple potential differences between ethnic groups – and to recognise that these may vary substantially across time, space and according to the groups under consideration. A full understanding of current patterns may also require consideration of historical experiences.

All ethnic contrasts must therefore be situated within a more general understanding of the current and past experiences of the groups in question. The remainder of this Chapter therefore summarises the migration history and current characteristics of the main ethnic groups in Britain.

3.2.1 Migration to Britain

Britain has long contained minority ethnic groups of Asian and African origin – for example, the nineteenth century saw a diverse collection of Indian sailors, domestic staff, princes and students [308]. Large-scale immigration, however, was a phenomenon of the second half of the twentieth century. This mass migration was facilitated by the 1948 British Nationality Act (offering British citizenship to anyone from the British colonies) and initially was primarily driven by the post-war British labour shortage. Migration slowed somewhat following the restrictions of the 1962 and 1968 Commonwealth Immigration Acts. Yet it by no means ceased, and continues to this day [295].

The earliest major migration wave was from the Caribbean (peaking in 1955-64), followed by migration from India and Pakistan (peaking in 1965-1974). During this period 134 000 Indians entered the UK. These were initially mainly single males staying temporarily, although the 1968 Immigration Act shifted the migrant profile towards families settling permanently. Immigration by East African Asians, mostly of Indian origin, also increased rapidly over this period. This was largely driven by the Africanisation programs in Kenya, Uganda, Tanzania and Malawi, peaking with the 28 000 individuals entering Britain in

1972 following the expulsion of Asians from Uganda [309]. This expulsion therefore triggered the unusual phenomenon of whole communities migrating together and with the intention of settling permanently. Bangladeshi immigration peaked in the 1980s and Black African migration in the 1980s and 1990s. The most recent arrivals have included many refugees and asylum seekers from a variety of areas, including the former Yugoslavia, Somalia and the Middle East. They have also included large numbers of economic migrants from recent EU members such as Poland.

In addition to this variation in migration timing, there has also been considerable variation in the characteristics of the migrants and the circumstances which they faced upon arrival. Indian immigrants contained an unusually high proportion of highly educated professionals, and this was particularly true of those migrating via East Africa [241, 309]. Yet both Indian and East African Asians were internally very diverse in terms of caste, religion and regional origins [310]. These different axes intersected with socio-economic differences – for example, as a group Hindus were more educationally and materially advantaged than either Sikhs or Muslims. These initial differences have in turn subsequently affected their experiences in Britain [295].

By contrast, Pakistani immigrants were more homogenous, primarily comprised of small-scale land-owning farmers from Pakistan [263]. They also showed greater homogeneity in their employment and geographical settlement, becoming heavily concentrated in the steel and textile industries in the Midlands and the North. The decline in these industries therefore had substantially greater adverse effects upon Pakistanis than Indians.

Thus despite migrating at similar times there were important differences in the characteristics and migration experiences within and between Indians and Pakistanis, and these help to explain the socio-economic differences seen today. The Bangladeshi immigration experience differed again in starting later and in peaking during a substantial downturn in the British economy. This contrasts with the earlier Caribbean, Indian and Pakistani migration waves, during which short-term improvements in the British economy were rapidly followed by increased immigration [295]. That Bangladeshis were an exception to this pattern helps to explain their particularly disadvantaged status today.

3.2.2 Current characteristics of Britain's ethnic groups

Population size

In 2001 the minority ethnic population of Great Britain was 4.6 million, forming 8.1% of the total population (Table 3.1). This included nearly a million children aged 5-15 (11.6% of all children of these ages), over 90% of whom will have been born in Britain [311]. Among adults the most common non-White ethnicity was Indian (1.8%), while in children Indian and Pakistani ethnicity were equally common (2.1%), and Mixed race ethnicity was the most common (2.7%).

Table 3.1: Ethnic composition of Great Britain in the UK census 2001

	All ages		**Age 5-15**	
	N	**Percent**	**N**	**Percent**
White				
• White British	50 366 497	88.2	6 992 057	86.4
• White Irish	691 232	1.2	28 411	0.4
• White Other	1 423 471	2.5	130 033	1.6
Mixed	673 798	1.2	218 159	2.7
Black or Black British				
• Black Caribbean	565 621	1.0	82 820	1.0
• Black African	484 783	0.8	94 724	1.2
• Black Other	97 198	0.2	24 970	0.3
Asian or Asian British				
• Indian	1 051 844	1.8	171 332	2.1
• Pakistani	746 619	1.3	172 317	2.1
• Bangladeshi	282 811	0.5	71 459	0.9
• Other Asian	247 470	0.4	39 423	0.5
Chinese	243 258	0.4	33 431	0.4
Other	229 325	0.4	30 650	0.4
Total non-White population	4 622 727	8.1	939 285	11.6
Total population	57 103 927	100.0	8 089 786	100

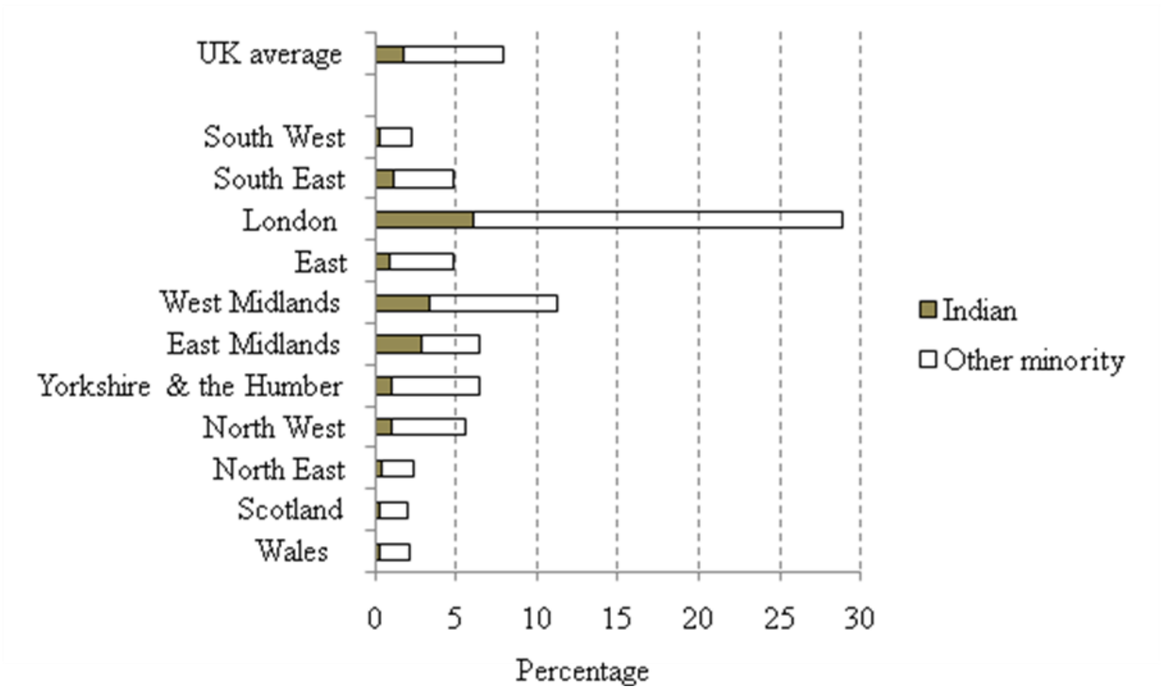
Sources: General registrar for Scotland 2002 [312], Office for National Statistics 2003 [313], Office for National Statistics 2004 [314].

Geographical distribution

Regional distribution

Twentieth century immigration to Britain was characterised by migrants acting as 'replacement populations' in areas where demand for labour was high but to which internal White migration was low. Thus relatively few migrants settled in areas of high unemployment like Wales or North East England, or in areas attractive to White workers like East Anglia. Instead, minority ethnic populations initially concentrated in the centres of large industrial cities such as London, Manchester and Birmingham. This was perpetuated as subsequent migrants headed preferentially for these areas, resulting in large regional differences today (Figure 3.1). Currently 48% of all minority ethnic individuals live in London and other major cities, and most of the remainder live in other metropolitan areas [261, 311].

Figure 3.1: Proportion of minority ethnic individuals by region in the 2001 UK census



Source: Office for National Statistics 2001 [315].

Area deprivation

Partly because of its concentration in urban areas, Britain's minority ethnic population is overrepresented in socio-economically deprived areas. More than half lives in the 12% most deprived local authority areas, these local authorities therefore containing over four times the average proportion of minority ethnic individuals [316].

Residential segregation and ethnic density

As well as being concentrated in Britain's cities, most minority ethnic groups show some within-city residential segregation – that is, clustering in particular areas as opposed to being distributed evenly through the population. Levels of residential segregation are, however, comparatively modest compared to countries like the US. In the 2001 census, 10-

20% of individuals from most minority ethnic groups lived in areas where over 87% of inhabitants were White British (20% for Indians), and a further 30-50% lived in areas which were between 50-87% White British (46% for Indians) [317]. Bangladeshis were the only group with over half of individuals (53%) living in areas less than 50% White British.

Bangladeshis are also unusual in having shown an increase in ward-level residential segregation in London since 1981 [295]. By contrast, Black Caribbean segregation has decreased steadily over this time as, more recently, has Indians segregation.

Cultural assimilation

Identity

The long-standing sociological interest in ethnic residential segregation has largely been driven by the hypothesis that this is both a marker and a driver of social assimilation; that is, political, socio-economic and cultural integration into wider society [318-319]. While this hypothesis has been problematised with regard to political and economic participation [320], it finds some support in the cultural assimilation of British minority ethnic groups. Thus Modood [262] argues that the Black Caribbean integration trajectory has been typified by mixing and ‘style-setting hybridity’ while South Asians have tended to form comparatively inward-looking ‘ethno-religious communities’. In line with their decreasing geographical segregation, this has recently become less true of Indians; Modood describes this as ‘waiting to assimilate’ until achieving middle classes status. I find this an intriguing characterisation, and in Chapter 12 I return to the possible importance of this integration strategy for understanding the Indian mental health advantage.

Inter-marriage

Nevertheless, Indians continue to resemble Pakistanis and Bangladeshis in having highly homogenous marriage patterns. Surveys from throughout the past two decades show that 91-99% of Indian, Pakistani and Bangladeshi men and women have marriage partners from the same ethnic group [295]. This suggests these groups have not abandoned a culturally pluralist strategy, and also indicates why most second- and third-generation South Asians remain able to classify their ethnicity as Indian, Pakistani or Bangladeshi.

Language

Finally, there are considerable differences between minority ethnic groups in their everyday use of English [241]. For example, although almost all White and Black Caribbean households in the Millennium Cohort Survey exclusively spoke English at home, this was true of only a quarter of Indian households (Table 3.2). The remainder spoke English and another language or, in 10.5% of households, another language only. Between 12% and 33% of Black African, Pakistani and Bangladeshi households likewise spoke no English in the home.

Table 3.2: Household languages in the Millennium Cohort Survey

	White (%)	Indian (%)	Pakistani (%)	Bangladeshi (%)	Black Caribbean (%)	Black African (%)
English only	97.5	26.3	9.4	2.7	97.4	45.4
English and other	2.0	63.3	66.9	64.2	2.5	42.6
Other only	0.5	10.5	23.7	33.1	0.1	12.0

Source: Panico *et al*.2007 [242].

Socio-economic position (SEP)

The forms of assimilation described in the previous section link to ethnic identity – that is, the meaningful collectivities to which individuals feel they belong. As highlighted previously in Section 3.1.1, it is also important to consider the social and structural basis of ethnicity. In migration sociology this is recognised as a distinction between behavioural assimilation and structural assimilation, and it has long been known that the one does not necessarily imply the other [321]. I therefore now turn to the structural assimilation of Britain's minority groups.

As a group, minority ethnicities fare worse on individual and household-level measures of SEP than the White British majority. Britain's minority ethnic populations are concentrated in lower occupational classes, and face higher unemployment, poorer working conditions, lower household incomes, poorer quality housing and higher household overcrowding [241, 316-317, 322-323]. Yet treating minority ethnicities 'as a group' masks considerable variation between groups and for different indicators of SEP. I therefore examine separately the profile of different ethnic groups for the four key SEP indicators measured in B-CAMHS: education; income; employment; and housing.

Educational attainment of adults

Table 3.3 presents data on educational attainment from the 2001 census. Pakistani and Bangladeshi adults are disadvantaged compared to White British adults, with more individuals with no qualifications and, in the case of Bangladeshis, fewer degree holders. Stratification by gender reveals that this applies to both sexes, but particularly to women [322]. Black Caribbeans have similar educational attainment to White British, and Black Africans, Indians and Chinese have similar or better attainment. For Indians and Chinese, it is also notable that adults age 16-74 contain more degree holders but similar numbers with no qualifications. This indicates a greater spread of educational attainment, as has also been observed in other surveys [324-325].

This may partly reflect a generational shift; as Table 3.3 shows, in 2001 Indians and Chinese aged 16-24 were both more likely to have degrees and less likely to have no qualifications. These differences are substantial, with the percentage of 16-24 year old Indian degree holders being almost twice that of Whites (19.6% vs. 10.4%). The Chinese percentage was even higher (25.1%) and the proportion was also somewhat larger in Black Africans (13.4%). These differences are particularly remarkable given that these groups are not advantaged over Whites for most other SEP indicators, and indeed often fare somewhat worse (see below). That education stands out in this way therefore suggests a cultural commitment to education which I argue in later Chapters may have important consequences for child mental health.

Table 3.3: Educational attainment by ethnic group, 2001

Ethnicity	**Age 16-74**		**Age 16-24**	
	Proportion with no qualifications (%)	**Proportion with degree-level or equivalent (%)**	**Proportion with no qualifications (%)**	**Proportion with degree-level or equivalent (%)**
White British	29.5	18.2	16.0	10.4
Black Caribbean	26.8	19.7	16.3	8.6
Black African	13.5	38.8	12.6	13.4
Indian	26.8	30.7	10.6	19.6
Pakistani	41.3	18.3	22.6	11.1
Bangladeshi	47.2	13.5	21.6	9.3
Chinese	25.6	37.3	8.9	25.1

Source: Census 2001, Table S117 [326].

Household income

Table 3.4 presents data from the Family Resources Survey 2000-2001 on the household income distribution of children aged 0-16. In all ethnic groups more than 20% of children live in the lowest income quintiles, this reflecting the overrepresentation of families with children among low income households. This trend is, however, particularly marked for minority ethnic groups. Most striking of all is the 70% of Pakistani and Bangladeshi children living in the lowest income quintile, as compared to 23% of Whites.

Table 3.4: Distribution of children aged under 16 across quintiles of household income by ethnic group, 2000-2001

Ethnic group of head of household	**Net equivalised disposable household income**				
	% in bottom quintile	**% in second quintile**	**% in third quintile**	**% in fourth quintile**	**% in top quintile**
White	23	24	22	18	14
Black Caribbean	28	30	16	14	12
Black Non-Caribbean	42	25	12	13	9
Indian	37	22	16	14	11
Pakistani/Bangladeshi	70	22	5	2	2
Other	41	17	13	13	16
All children	25	24	21	17	13

Source: Department for Work and Pensions 2002 [327, table 4.1]. Income equivalised for couple status and number of children.

Employment

The 2001 census revealed marked variation in the labour market profiles of different ethnic groups [317]. Pakistani, Bangladeshi, Black Caribbean and Black African men all faced substantial disadvantage, with the highest rates of economic inactivity and unemployment (12-16% unemployed vs. 5% in the White British). Moreover, many economically active Bangladeshi and Pakistani men were employed part time (44% Bangladeshi and 18% Pakistani vs. 5% White British). By contrast the profile of Indian and Chinese men was broadly similar to that of White British men. For example, 82% of Indian men were economically active (vs. 83% White British), of whom 7% were working part-time (vs. 5% White British). Indian women were less likely to be economically active than White British women (62% vs. 71%), but those who were economically active were more likely to work full-time (69% vs. 56%).

Yet while achieving similar labour market outcomes to White British adults in absolute terms, Indian and Chinese adults nonetheless do less well than would be predicted by their high educational attainment [317]. Thus at any given level of qualifications, Indian and

Chinese adults have lower economic activity and higher unemployment than their White British counterparts. This ethnic penalty is also observed for most employment outcomes in the other main minority groups. This may partly reflect the substantial downward social mobility which many professionals experienced upon migrating to Britain [241]. This is not the whole explanation, however, as the ethnic penalty is little changed when restricting analyses to those born in Britain [317].

Housing tenure

Indians and Pakistanis have long shown a strong preference for home-ownership [295]. In the 2001 census, owner occupation was highest in Indians (76%), followed by White British (70%) and Pakistanis (67%) [328]. Yet despite similar rates of home-ownership, Indians and Pakistanis differed considerably in the nature of their housing stock: many Indians owned detached and semi-detached houses in the suburbs, while most Pakistanis owned inner-city terraces [295]. Indian households were also less likely than Pakistani households to be over-crowded (18% vs. 26%), although still substantially more likely than White British (6%). Indeed, over-crowding was much more common in all minority ethnic groups compared to White British, with rates of 15-30% in most minority groups and over 40% in Black Africans and Bangladeshis. Black Africans and Bangladeshis were also the groups with the lowest rates of homeownership (26% for Black Africans and 19% for Bangladeshis), with around half of both groups living in socially rented accommodation.

Ethnic heterogeneity in SEP profiles

To summarise, the overall pattern of disadvantage conceals considerable heterogeneity between ethnic groups. For many SEP indicators, Indians and Chinese show broad parity with Whites, Black Caribbeans and Black Africans are more disadvantaged, and Pakistanis and Bangladeshis (particularly the latter) fare worst of all. Yet this pattern does not apply to all measures of SEP; for example, Black Africans are not disadvantaged in terms of education, and Pakistanis are not disadvantaged in terms of housing tenure. This reflects the important point of diversity between ethnic groups in the *inter-relationship* between SEP indicators.

The labour market ethnic penalty described above provides one example of this, demonstrating that at a given educational level minority ethnic individuals have worse employment outcomes. Conversely, within any given occupational class, minority ethnic members have better than average education [241]. Disparities across ethnic groups have also been reported for other SEP indicators [241, 303-304, 307, 329]. For instance, Table 3.5 presents data from the Fourth National Survey of Ethnic Minorities. All minority ethnic groups had markedly lower incomes than Whites in the same occupational social class – for example only half the White average for Pakistanis and Bangladeshis. Likewise, length of unemployment was far longer among unemployed minority ethnic men, and quality of accommodation was far lower among minority ethnic homeowners. Although particularly large for Pakistanis and Bangladeshis, these multiple ethnic penalties were also observed for Indians.

Table 3.5: Ethnic variations within socio-economic bands in the Fourth National Survey of Ethnic Minorities

	White	**Black-Caribbean**	**Indian & African Asian**	**Pakistani & Bangladeshi**
Mean income (£) within Registrar General's occupational social class				
I/II	250	210	210	125
IIIN	185	145	135	95
IIIM	160	145	120	70
IV/V	130	120	110	65
Unweighted base	*1894*	*869*	*1142*	*969*
Mean duration unemployment (months) among unemployed	7	21	12	24
Unweighted base	*128*	*91*	*91*	*166*
Percent lacking one or more basic housing amenities†, by housing tenure				
Owner occupied	11	12	14	38
Renters	27	23	28	37
Unweighted base	*2867*	*1205*	*2001*	*1776*

Source: Nazroo 1997 [307, p.99].
† Corresponding to exclusive use of: bath or shower; bathroom; inside toilet; kitchen; hot water from a tap; and central heating.

Apparently similar socio-economic indices may therefore mean different things for different ethnic groups. This applies not only to comparisons with Whites but also to comparisons between different minority groups, as illustrated by the differences in housing stock between Indian and Pakistani home-owners. As discussed in Section 3.1.3 p.73, single indicators of SEP may over- or under-adjust for other facets of SEP when making ethnic comparisons. This may, in turn, create a misleading impression of how far SEP explains observed ethnic differences. The consequent need to use multiple SEP indicators informs my analysis approach in Chapter 10 and Chapter 11.

Family composition – household size and household structure

South Asian household sizes are substantially larger than those of other ethnic groups. In 2002, average household size was 4.7 people in Bangladeshis, 4.2 in Pakistanis and 3.3 in Indians. By contrast, the average was under three in all other ethnic groups, being 2.3 in Whites [261]. Part of the reason for this larger South Asian household size is a higher prevalence of three-generation families and a far lower prevalence of lone parent families. For example, in 2002 the prevalence of lone parent families among families with dependent children was over 50% in households headed by Black Caribbean or Mixed race individuals, but only 9% for Indians and 15% for Pakistanis. Whites were intermediate, with a prevalence of 23% [261].

Educational achievement of children

The educational achievement of children at school shows the same striking pattern of Indian and Chinese advantage as is seen for young adults with respect to completion of higher education. At GCSE level, Indian children have out-performed White British children since at least the early 1990s and this advantage has increased over time [330]. The result is that in recent years Indian students have performed substantially better than their White British counterparts, as have Chinese students. For example, as shown in Table 3.6, in 2004 the proportion of Indian boys achieving five or more good GCSEs was 61.6% vs. 47.3% in White British boys. In girls the corresponding proportions were 71.9% and 57.3%. By contrast, Black Caribbean pupils achieved worse GCSE results than White British students as, to a lesser extent, did Black African, Pakistani and Bangladeshi pupils.

Table 3.6: Proportion of pupils in England achieving five or more A*-C GCSE grades/GNVQs: by ethnic group and sex, 2004

	Boys (%)	Girls (%)
White		
• White British	47.3	57.3
• White Irish	54.0	62.5
Mixed	44.8	54.4
Black or Black British		
• Black Caribbean	27.3	43.8
• Black African	37.3	48.9
• Black Other	29.8	43.0
Asian or Asian British		
• Indian	61.6	71.9
• Pakistani	38.8	52.1
• Bangladeshi	41.0	55.2
Chinese	69.5	79.1
Other	43.0	54.4
All pupils	46.8	57.0

Source: Department for Education and Skills 2005 [331].

The pattern in Table 3.6 is observed across the school years and for other educational indicators. In 2002, Indian and Chinese children did much better than White British in Key Stage tests at every age, while Black Caribbean, Black African, Pakistani and Bangladeshi pupils did worse [330]. SEP seemed to account for some but not all of these differences. Indian, Chinese and White children were also less likely to be recorded as having special educational needs than Black Caribbean, Black African, Pakistani and Bangladeshi children. More recent data confirms the higher attainment of Indian and Chinese students, and also reveals that they make better *progress* between Key Stages and that their absolute advantage therefore widens with age [332].

In Chapter 12 I discuss the possible importance of this substantial advantage for understanding the Indian mental health advantage. Unfortunately, however, the origin of this high academic achievement has to date received little attention from educational researchers. As such, little is known about how far it reflects a cycle of educational advantage stemming from the higher educational attainment of Indian and Chinese adults. Similarly, little research has examined the contribution or mediating influence of factors like parental aspirations or engagement in their child's education. Several qualitative studies indicate that Indian and Chinese parents place a high value on education and actively support their children [333-335], but interpreting these findings is complicated by the absence of comparison groups. In one study which did include a comparison group, however, there was intriguing evidence that the nature of parents' aspirations differed between 'South Asians' (mostly Indians and Pakistanis) and Whites. Specifically, South Asian parents stressed success in school as a route to greater confidence and self-advancement, while White parents more often said they wanted their child to do as well as they could and to enjoy school [214]

Larger quantitative surveys by the Department for Children Schools and Families have likewise not yet realised their potential to shed light on the causes of the Indian and Chinese educational advantage. This is because these surveys have either used the Indian and Chinese advantage as a reason to exclude them from minority ethnic over-sampling [336] or else have used only the meta-ethnic group of 'Asian/Asian British' [337]. This latter study found Asian and White parents were equally likely to report that they were 'very involved' in their child's education. They also had similar views on parental responsibility for their child's education and on the importance of ensuring regular school attendance. Asian parents were, however, more likely to think it was 'extremely important' to help their child with homework (82% vs. 72%) and to report doing this most of the time (69% vs. 56%). Asian parents were also more likely to be involved in school-related activities such as homework clubs or Parent Councils. These findings are therefore suggestive but inconclusive, being limited by a failure to disaggregate the Asian groups and to investigate the causes of ethnic advantage as well as disadvantage. These limitations parallel those discussed in Section 3.1 for the child mental health literature.

Physical health and disability

The health of Britain's minority ethnic adults shows some heterogeneity with regards to outcome considered, but frequently shows a pattern similar to that outlined for socio-economic indicators. Data from the 2001 census indicates that Pakistanis and Bangladeshis had by far the worst self-rated health, particularly in females [338]. Indian and Black Caribbean self-reported health was closer to that of Whites, although still somewhat worse. Black African and Chinese individuals had similar or better health than the White population. Other population-based surveys report a similar pattern, both for self-reported health and for other health indicators including limiting long-term illness, being registered disabled or economic inactivity due to permanent sickness/disability [261, 307, 316-317, 339-340].

In school-age children, much research on ethnic differences in child health focuses on specific conditions such as Vitamin D deficiency [340] rather than physical health in general. One important exception is the 1999 Health Survey for England, which included over-sampling from minority ethnic groups [341]. It found that health outcomes showed striking variation in their pattern of inter-ethnic differences. As Table 3.7 shows, for self-reported health the pattern in children is similar to that in adults, with most minority groups reporting somewhat worse outcomes than the general population (i.e. majority White British). The reverse is true, however, for limiting longstanding illness and acute recent illness, which are the same or less common in all minority ethnic groups. Similar discrepancies across different health indicators were also reported in the British General Household Survey (1991-94) [342] and in the London sample of the 1991 census [342]. For example, the British General Household Survey found that Indian, Pakistani and Bangladeshi children had higher rates of GP consultation than White British, but simultaneously had lower rates of hospital usage [339].

Table 3.7: Inter-ethnic differences in health status for children aged 2-15 in the Health Survey for England, 1999

		General population (%)	**Black-Caribbean (%)**	**Indian (%)**	**Pakistani (%)**	**Bangladeshi (%)**	**Chinese (%)**
Poor self-reported general health	**Boys**	9	12	11	13	16	9
	Girls	8	10	13	8	11	9
Limiting longstanding illness	**Boys**	11	9	10	10	7	8
	Girls	9	9	6	7	7	4
Acute sickness in past two weeks	**Boys**	14	10	9	10	5	6
	Girls	14	11	6	7	6	5

Source: Nazroo *et al.* [343].
Data obtained by parent report for children aged 12 or under.

The source of these discrepancies between health outcomes is unclear, and interpretation is complicated by the reliance upon reported health and service use. Reported health is problematic because individuals will inevitably vary in their response thresholds, particularly in questions which leave substantial room for individual interpretation (e.g. “Overall, how is the health of your child?”). There is therefore the same potential for systematic reporting differences between groups as exists for mental health assessments (Section 2.4.2, Chapter 2). Thresholds may also differ for health seeking behaviour, which may likewise lead to systematic group differences in service use.

Yet unlike mental health, objective measures do exist for many aspects of physical health. Unfortunately, few studies to date have used such measures. One exception is the use of detailed anthropometric measures and blood tests to demonstrate that British South Asian children show the same tendency to insulin resistance which is observed in adults [344]. The Millennium Cohort Study should also prove highly informative, particularly through its coverage of multiple health outcomes and its triangulation of objective and subjective measures [345]. Although findings thus far largely relate to infant development, a few reports of older children have now been published. This includes the finding of lower rates of wheeze and asthma in Bangladeshi three year olds [242], thus again indicating a disjunction between ethnic patterns in childhood health and what one might expect based on adult health and SEP.

Finally, some studies have studied inter-ethnic differences in health risk behaviours. One of the largest is the RELACHS study, which investigated substance use among children aged 11-14 in London. It provided strong evidence that, compared to White British, regular smoking and regular drinking were less common in Indians, Pakistanis, Bangladeshis and Black Africans [346]. Indeed, in all minority ethnic groups the direction of the effect was that these risky behaviours were less common than in White British.

Experiences of racism

Racism and racial harassment remain a common experience for minority ethnic individuals. In the mid 1990s approximately one in eight minority ethnic individuals reported at least one incident of racial harassment in the past year [241]. The situation may recently have improved somewhat – the British Crime Surveys suggest a decreasing incidence of racially motivated attacks over the 1990s, falling from an estimated 390 000 attacks in 1995 to 280 000 in 1999 [261]. Nevertheless, racially motivated attacks still constituted 12% of all crime against minority ethnic individuals, as compared to 2% for Whites. Moreover, such attacks are particularly distressing, with 42% of victims reporting themselves to be ‘very much affected’ as compared to an average of 19% for all crimes [261].

Yet while racism has always been a problem for visible minority groups in Britain, there is some evidence of changes in recent times in the groups most strongly affected. In the past, minority ethnic groups which were more culturally self-sufficient may have been buffered against the experience of societal racism. There is consistent evidence that over many decades of the late twentieth century White Britons displayed similar prejudice towards all visible minorities, but that Black Caribbeans were more aware of this discrimination [347]. This is plausibly because the more assimilatory behaviour of Black Caribbeans led them to have more contact with wider British society.

More recently, however, lower cultural assimilation may carry risks of its own. Both official crime statistics [261] and attitude surveys [241, 348] suggest that today it is South Asians who face the most prejudice. This may reflect a new cultural racism which denies a colour-based prejudice but which affirms the naturalness of wanting to live with ‘people like us’ [349]. These attitudes are particularly directed towards Pakistanis and Bangladeshis, and have been sharpened by the recent rise of Islamophobia in the West.

3.3 Chapter summary and conclusions

Chapter 2 emphasised that when comparing mental health across ethnic groups, one must build a case for the validity of one’s classification system and investigate the possibility of systematic bias in one’s measures. Where ethnic groups do genuinely differ in their mental health, I believe that studying these may be of public health value. Where inequalities exist, describing these is usually necessary in order to take steps to reduce them. Investigating ethnic differences may also generate insights regarding the underlying aetiology and causal mechanisms – epidemiology is, after all, fuelled by analysis of health differences between populations.

Yet to generate such insights, observed ethnic differences must be used as a starting point for further analyses. British minority ethnic groups differ from each other and from the White majority in multiple ways. These include differences in geographical distribution and area deprivation; socio-economic position; family composition and family size; academic attainment; physical health and substance use; and exposure to acute and chronic stressors through racism. Many of these are correlates of child mental health problems (see Section 2.2.4, Chapter 2) and therefore plausible mediators or confounders.

Direct investigation is therefore needed to establish which factor or factors are important in explaining any observed ethnic difference. This requires direct measurement of potential mediators and confounders, and attention to the complexity of adjusting for confounders such as SEP. Ethnic comparisons are also especially likely to be productive if consideration is given to the potential for variation within meta-ethnic groups such as 'Black' or 'South Asian', and if groups with better health outcomes are studied alongside groups with worse outcomes. In the next Chapter I examine how far the existing literature achieves these key objectives by presenting a systematic review of studies of ethnicity and child mental health in Britain.

Chapter 4 Ethnicity and child mental health in Britain: systematic literature review

4.1.1 Introduction

Evidence on the mental health of children from different ethnic groups in the UK is scattered, with previous reviews being non-systematic and sometimes focussing only on particular ethnic groups or particular types of mental health problem [174, 350-363]. In this Chapter I present the first systematic review of this topic, synthesising evidence from population- and clinic-based studies in Britain published over the last 40 years on the major child mental health problems in all minority ethnic groups.

I chose a broad scope for my systematic review in order to provide a context for my findings on Indians. I therefore outline the findings of my systematic review in full, although focusing on population-based studies of common mental health problems which contain an Indian sample.[9] I also provide further context for my findings by presenting a non-systematic review of the main population-based studies of common mental disorders in adults in Britain, and child mental health problems in India.

4.1.2 Methods

<u>Review objectives</u>

This review was motivated by two questions:

1. How, in population-based studies sampling from the general population, does the prevalence and proportional morbidity of mental health problems differ among children from different ethnic groups in Britain?
2. How do ethnic differences in levels and patterns of service use observed in clinic-based studies compare with estimates of prevalence and proportional morbidity obtained from population-based studies?

---

[9] For the published version of this review, which includes a fuller discussion of other ethnic groups, less common disorders and the findings of clinic-based studies, see 364. Goodman, A., V. Patel, and D.A. Leon, *Child mental health differences amongst ethnic groups in Britain: a systematic review*. BMC Public Health, 2008. **8**(1): p. 258., Appendix 3.

Search strategy

I aimed to identify all relevant quantitative studies produced at any time up to and including June 2007, following the guidelines of the expert working group consensus statement on the Meta-analysis of Observational Studies in Epidemiology (MOOSE)[365].

Between January and July 2007 I searched keywords, titles and abstracts in 16 electronic databases and eight websites (see Box 14.1, Section 14.1 Appendix 2). The search string combined a wide range of free text terms and subject index headings, and was evaluated and refined by assessing retrieval of known studies. I scanned reference lists of previous discussions of the literature and non-systematic reviews [174, 350-363], and also of all articles considered for inclusion. Studies eligible for inclusion in this review were entered into the Science Citation Index to identify studies which had cited them. To locate other relevant work, particularly unpublished studies, I asked for suggestions from experienced researchers in the field, circulated requests for assistance to five special interest groups (see Box 14.1), and contacted the corresponding authors of studies eligible for inclusion and published in the past 20 years. 35 first/corresponding authors could be traced and were contacted in July 2007. Finally, I sought to locate large epidemiological population-based studies of child mental health in Britain, because these seemed particularly likely to contain relevant information which would not necessarily be reported in an abstract. I located these through existing reviews [4, 201, 239, 366-367] and by consulting other researchers.

Inclusion criteria

My inclusion criteria were as follows:

- **Participants**: Living in Britain; aged 0-19 years; sampled from the general population or from mental health clinics serving the general population (i.e. not small and selected groups such as foster children or children in secure forensic units).
- **Ethnicity**: I operationalised ethnicity to include groups as defined by the 2001 UK Census [368]. Additional categories were added to cover groups whose religion, language or way of life serves in Britain as a marker for membership of a particular 'meaningful collectivity'. This included groups such as Orthodox Jews and Travellers but not, in the absence of additional information, internally diverse

groups such as Christians or Muslims. Minority groups defined simply as 'minority', 'non-White' or 'other' were excluded. Included studies had to contain 1) at least two specified ethnic groups (not necessarily with one White/White British), or 2) one minority group compared to all other children in the sample/a 'comparable' general population sample (see Section 14.1 Appendix 2, p.401).

- **Mental health**: Included outcomes were: referral or admission to a child mental health service; "a psychiatric diagnosis" (unspecified) made by a mental health specialist; emotional disorders; behavioural disorders; hyperactivity disorders; less common disorders, including psychosis, autistic spectrum disorders and eating disorders; somatoform disorders; suicide and deliberate self-harm (DSH). Only validated clinical interviews or questionnaires were accepted, but validation in each ethnic group was not required. An experienced psychiatric epidemiologist judged whether enough evidence existed to establish the validity of interviews and questionnaires, doing so blind to study findings.
- **Study types**: Included study types were: 1) Population-based studies of prevalence or mean scores (minimum sample size N≥40 for each included ethnic group for prevalence, N≥10 for each included ethnic group for mean scores); 2) Clinic-based studies of the relative proportion of referrals/in-patients in clinics from ethnic minority groups, as judged against the ethnic composition of a base population such as the local catchment area (no minimum sample size); 3) Clinic-based studies which compared ethnic groups in terms of their proportional morbidity from different diagnoses – that is, the relative frequency of emotional disorders, of behavioural disorders etc among all mental health diagnoses. (minimum sample size N≥20 for each included ethnic group).
- **Minimum sample sizes**: The minimum sample sizes described above were imposed to avoid highly underpowered studies leading to 'uninformative' null findings and/or publication bias. They varied for different study types depending on the estimated power to detect effects (see Section 14.1 Appendix 2, p.401).
- No restriction was made on date or language of publication.

Assessing studies for inclusion, data extraction and data analysis

Handling abstracts and assessing studies for inclusion.

I assessed all titles and abstracts (N=6286) for possible relevance. A test-retest evaluation 4 weeks apart on 1391 of the electronically-retrieved studies demonstrated good reliability in this; I re-identified 42 of the original 43 studies and no additional papers. Studies judged as potentially within the scope of the review were independently assessed for inclusion by me and by a second epidemiologist, with disagreement decided by consensus. When assessing studies for inclusion and when subsequently extracting data from them, I attempted to contact authors whenever important details were unclear.

Extracting data and assessing methodological limitations

I extracted data according to pre-determined fields for all mental health outcomes and all ethnic groups meeting my inclusion criteria. I also judged studies against a predetermined list of possible methodological limitations devised for the purposes of this review. These included limitations in the measurement of mental health; limitations in the measurement or reporting of ethnicity; methodological limitations which could cause selection or information bias; and the potential for confounding by age, sex and socio-economic position (see Box 14.2, Appendix 2 p.403). Data extraction and assessment of limitations were independently checked by a second epidemiologist, with the rare instances of disagreement decided by consensus. One paper was published in Spanish [369], and my data extraction was independently reviewed by a native Spanish speaker.

In some studies (indicated in Table 14.4 and Table 14.5 in Section 14.1 Appendix 2) the relevant statistical tests were not reported but were 1) calculated by me using data in the paper 2) calculated by me using data provided by the authors or 3) provided for me through further data analysis by the study authors. These calculations involved simple statistical tests such as t-tests or one-way ANOVAs to compare the means between particular ethnic groups, or chi-squared tests to compare the relative frequency of different types of problem.

Data analysis

I judged that a formal quantitative meta-analysis was impossible because the classifications of ethnicity and of mental health outcomes were too heterogeneous. Instead I adopted a semi-quantitative descriptive approach which categorised the results of individual analyses according to whether each minority group showed evidence of more or fewer mental health problems than the White/White British/general population children in the study (for other examples of similar approaches see [370-371]). Combined categories were used for studies where children from the same ethnic group showed discrepant findings by mental health outcome, informant or gender. For example, if boys from a particular group showed an advantage but girls were no different, the combined category 'fewer problems/no difference' was used.

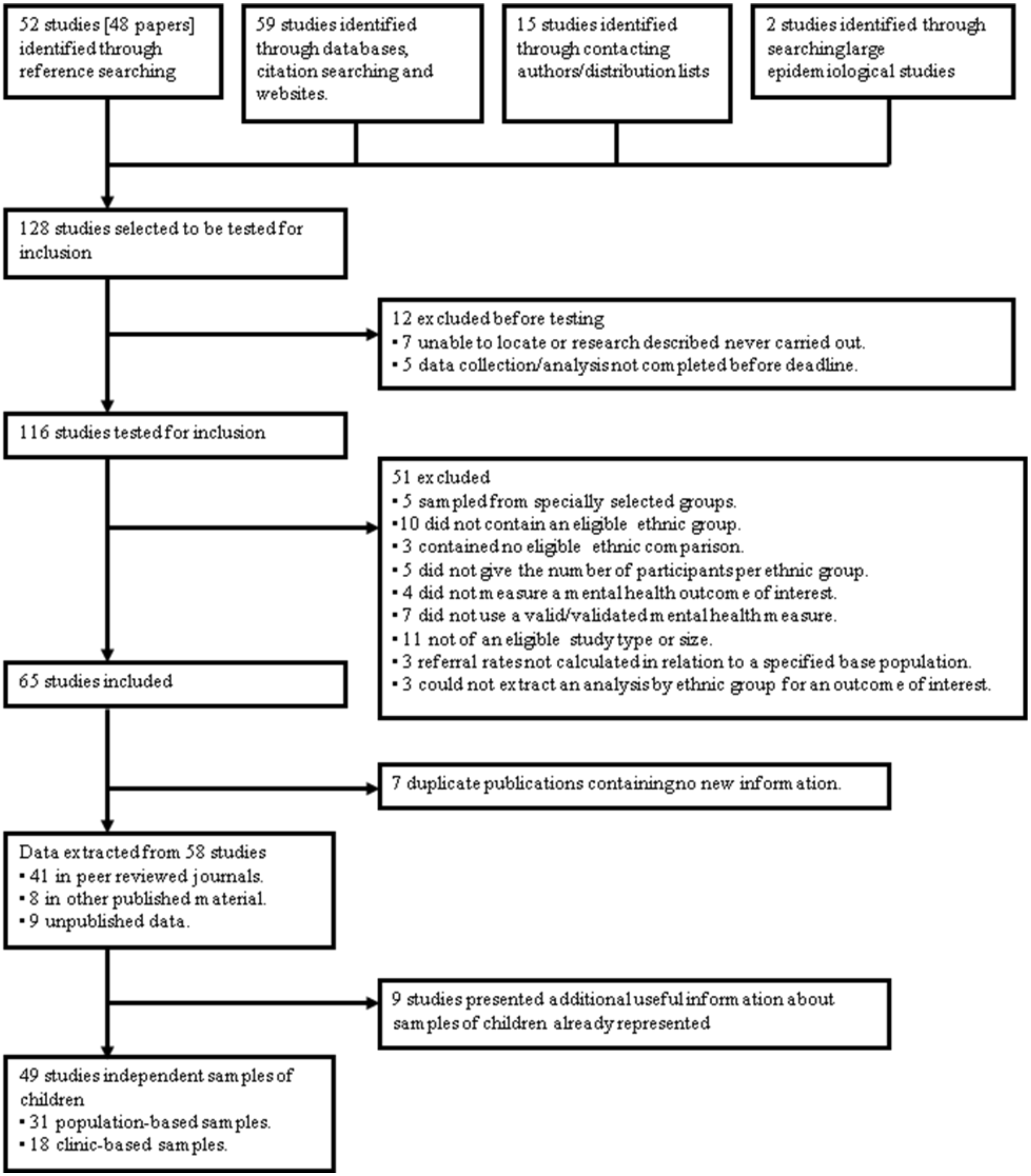
In three studies [214, 372-373], containing eight minority ethnic study populations, ANOVA analyses provided evidence ($p<0.05$) of differences in mean scores, but post hoc contrasts between specific groups were not presented. These eight study populations could therefore only be tentatively grouped, based on the trend showed by the mean scores. Some studies presented not only the raw comparisons but also models which adjusted for a range of potential confounders. In such cases, I used the results of models adjusting only for age and gender or, failing that, I used the raw data/unadjusted models. This was done in accordance with my primary aim of *describing*, rather than explaining, ethnic differences.

4.1.3 Results

Study retrieval

Search results of electronic database, citation and website searching are shown in Table 14.1 (Section 14.1 Appendix 2), and Figure 4.1 summarises the origins of all studies. in total, 128 studies reported in 125 potentially relevant papers were identified, of which 116 studies had been completed and were successfully retrieved. Of these, 58 of these studies were excluded (for details, see Table 14.2, Appendix 2) and 58 were included. The 58 included studies covered 49 independent samples of children, of which 31 were population-based [2-3, 214, 238, 343, 372-394] and 18 clinic-based [395-412]. Of the 49 independent studies, 13 population-based studies contained an Indian subsample [2-3, 343, 373, 375, 377, 384, 387-388, 391-393, 413] as did one clinic-based study [405]. Nine further population-based studies presented additional informative information on samples of children already represented [157, 414-421]. These included studies are described in detail in Table 14.4 and Table 14.5 (Section 14.1 Appendix 2).

Figure 4.1: Selection of studies into systematic review



Description of studies

The 49 independent studies included in this review were predominantly recent (25/49 published since 2000, or 10/14 for studies with an Indian sample), reflecting the increasing interest in the mental health of children in minority ethnic groups. The studies were reasonably well-balanced with regard to the age of the children they contain, but were mostly conducted in England (45/49 studies). Of the 14 studies containing Indian samples, the two B-CAMHS surveys took nationally representative samples from Britain. Two further studies took nationally representative samples from England, nine sampled from large English cities (four from London, five from industrial Northern cities) and one sampled from a suburban area of South England. For a full summary of the characteristics of the the 49 independent studies included in this review, and the 14 stidies with an Indian sub-sample, see Table 14.3, Section 14.1 Appendix 2.

Of the 31 population-based studies, 23 reported 'all disorders' and/or a common child mental disorder; seven reported disordered eating attitudes; and one reported psychotic-like experiences. 13 of these studies included Indian samples, 12 reporting all disorders and/or a common child mental disorder and one reporting disordered eating attitudes. Of the 18 clinic-based studies, 15 examined over- or underrepresentation of ethnic groups relative to the base population (including one Indian sample), and seven examined proportional morbidity from different disorders (including zero Indian samples).

Although not specified in my inclusion criteria, all 49 studies included a White/White British/'general population' (i.e. largely White British) sample. This permitted a single strategy for combining information across studies, by always comparing the results for each minority ethnic group to the White/White British sample. Of the minority groups listed in the UK census, only Black Caribbean, Indian and Bangladeshi children were included in ten or more studies, while White Minority and Chinese children featured in five or fewer.

All studies had some methodological limitations, the most common being limitations in the measurement or reporting of ethnicity (42/49 studies, or 9/14 for studies containing Indian sample); the measurement or reporting of SEP (37/49, or 6/14 for studies containing Indian samples); and potential selection bias through clinic-based sampling and/or low response rates (30/49 or 4/14 for studies containing Indian samples). In addition, 10 of the 13 population-based studies which contained Indian samples had only questionnaire measures of mental health. This included 10/11 of the non-B-CAMHS studies.

Population-based studies of common child mental health problems

Table 4.1 summarises the results of population-based studies for the common child mental health problems. Black African and Indian children appeared to enjoy better mental health than White British children, with at least one finding of an advantage reported in 5/6 studies of Black Africans and 8/12 studies of Indians. This included 6/10 non-B-CAMHS studies of Indians. By contrast, most studies of Black Caribbean, Pakistani and Bangladeshi children found no evidence that their mental health differed from that of White British children. The mental health of Mixed race children also appeared similar to that of White British children, although the diversity of this group complicates interpretation of this finding. Similarly the inconsistent findings from studies of 'Black' or 'South Asian' children are hard to interpret given the potential heterogeneity of these ethnic categories. For other ethnic groups, including White minority or Chinese children, there was insufficient evidence to draw conclusions.

Table 4.1: Summary of findings of population-based studies of common child mental health problems

Ethnic group	No. study populations		Mental health problems/disorders relative to White/White British/'general population' children				
			Fewer problems	Fewer/no difference	No difference	More/no difference	More problems
White Irish	2		0	0	1	1	0
White minority (unspecified)	2		0	0	1	1	0
Mixed race	5		0	0	4 (?+1)	0	0
Black Caribbean	11		0	1	6	1 (?+1)	1 (?+1)
Black African	6 (in 5 papers)		3	2	1	0	0
'Black'	4		1	0	2	0	0 (?+1)
Indian	12		7 (?+1)	0	2	2	0
Pakistani	6		0 (?+1)	0	4	1	0
Bangladeshi	6		0	1	5	0	0
'South Asian'	5		1	1 (?+1)	1	0 (?+1)	0
Chinese	2		0	1	1	0	0
Orthodox Jewish	1		0	1	0	0	0

All differences significant at the 5% level, except for those shown in parentheses (e.g. '(?+1)') where the significance level for the specific contrast was not reported and where the study is therefore grouped by its trend. Not all studies are independent, as some compare children from several minority ethnic groups to the same White 'reference' group.

For most ethnic groups I was unable to identify study characteristics which might explain discrepant findings between different studies. In Indians, the only factor which seemed potentially important was the type of common mental health problem studied. Seven studies, containing 12 study populations, distinguished between emotional, behavioural and, in most cases, hyperactivity problems. This included five study populations of Indian children. These consistently found that overall Indian advantages reflected fewer behavioural/hyperactivity problems (3/3 studies [3, 377, 391] including B-CAMHS04), while overall Indian disadvantages resulted from more emotional problems (2/2 studies [343, 387]). The converse was suggested by four studies of Black Caribbean or Mixed White/Black Caribbean children [238, 394, 414], which indicated relatively more behavioural problems (3/4 study populations, although in one case in girls only [238, 414]) and/or relatively fewer emotional problems (2/3 study populations [414]).

Few studies containing Indian samples presented results disaggregated by gender, and the four which did were inconsistent. One reported an Indian advantage which was entirely attributable to better mental health among boys [391], while another found poorer mental health only among girls [343]. The remaining two studies found no evidence of a gender-ethnicity interaction [373, 387].

Overview of evidence for other types of studies

In this PhD I focus upon the common mental disorders. I therefore provide only a brief summary below of my review's findings for less common disorders and clinic populations, particularly as only two of these studies included Indians samples. For a fuller account, see [364] (reproduced in Appendix 3).

Population-based studies of less common disorders

Population-based studies provided consistent evidence that 'South Asian' girls scored higher on questionnaire measures of disordered eating attitudes (6/7 studies) [378, 381-382, 384-386], but not [380]). Only one study used clinician-based diagnoses in addition to questionnaires, and reported some evidence of a higher prevalence in South Asian girls [378]. Only one study disaggregated South Asians, and reported that higher scores were confined to Bangladeshi girls, with Indians and Pakistanis scoring similarly to Whites [384]. There was no evidence of a difference from Whites in the two small Black populations [385-386] but some evidence of an excess in the one, small Mixed race sample [385].

One population-based survey investigated psychotic-like experiences, reporting higher rates in Black Caribbean children, lower rates in South Asian/Chinese children and no evidence of a difference for White minority, Black African or 'Other' ethnicity children [422].

Clinic-based studies of proportional representation of ethnic groups in clinic populations

Table 4.2 summarises the results of clinic-based studies of proportional representation of ethnic groups among clinic populations. The seven study populations of Black Caribbean, Black African or ‘Black’ children were small, suffered serious methodological limitations, and produced inconsistent findings. By contrast there was far more consistent evidence of underrepresentation of Indian, Pakistani, Bangladeshi and ‘South Asian’ children, this being seen in 10/13 study populations. This included the one study which disaggregated South Asians and showed an underrepresented of Indian children [405].

Table 4.2: Summary of findings of clinic-based studies of proportional representation of ethnic groups among clinic populations

Ethnic group	**No. study populations**		**Proportional representation relative to the base population**		
			Under-represented	**Represented as expected**	**Over-represented**
Black Caribbean	1		0	1	0
Black African	1		0	1	0
‘Black’	5		2	1	2
Indian	1		1	0	0
Pakistani	2		2	0	0
Bangladeshi	4		4	0	0
‘South Asian’	6		3	2	1

All differences significant at the 5% level

Clinic-based studies of proportional morbidity of different disorder types

Six clinic-based studies examined proportional morbidity for common mental disorders, none of which contained an Indian sample (for details see [364], reproduced in Appendix 3).

4.1.4 Broader context

Adult mental health of minority ethnic groups in Britain

Ethnic differences in adult mental may be doubly relevant to understanding ethnic differences in child mental health. First, parent mental health problems are strongly associated with child mental health problems (see Section 2.2.4, Chapter 2). Second, comparing child and adult mental health may provide insights into when in the lifecourse mental health problems develop. Unfortunately, adult mental health surveys rarely cover behavioural or hyperactivity symptoms. Instead ‘common mental disorders’ (CMD) in adults refer specifically to depression and anxiety, and are therefore most comparable to emotional problems/disorders in children.

In the past 15 years there have been three large, nationally-representative surveys which assessed CMD in minority ethnic adults in Britain, all using a fully structured clinical interview (the Revised Clinical Interview Schedule [423]). The EMPIRIC survey [244] and National Survey of Psychiatric Morbidity [424] found little evidence of inter-ethnic differences in CMD, while the Fourth National survey reports differences which are somewhat more pronounced but still not dramatic [243, 307] (for full results, see Table 14.6, Section 14.1, Appendix 2). The three studies also produced inconsistent findings on the relative advantage/disadvantage of particular minority groups. This includes Indian adults: the Fourth National Survey suggested fewer CMD while in EMPIRIC there was no evidence of an overall difference but a disadvantage in women aged 55-74 (the National Survey of Psychiatric Morbidity used only the meta-ethnic group ‘South Asian’).

Interpreting these findings is complicated by the possibility that, as for child mental health, systematic differences may exist across groups in mental health reporting. This also complicates interpreting one of the few consistencies between the studies, namely the finding in both the Fourth National Survey and EMPIRIC that substantially lower rates of CMD among Bangladeshis were largely confined to those who had recently arrived in the UK or were not interviewed in English.

Child mental health in India

In most countries Indians are a small fraction of the minority ethnic population. This includes major destination countries such as United States, Canada or Australia, and I have been unable to locate any large population-based surveys from these countries which present results separately for Indians. I therefore focus instead on results from India itself.

That child mental health is a key issue in India is clear from the exceptionally high suicide rates among teenagers, particular females, in Southern India [425]. For the purposes of this thesis, however, it is the prevalence of common mental health problems which is of greatest interest. This question has been addressed by three large, population-based surveys from India in the past 20 years [426-428], as summarised in Table 4.3.[10] For comparison, figures from the (pooled) B-CAMHS surveys are also presented.

Table 4.3: Population-based, cross-sectional studies of common child mental health disorders in India

Setting, date and reference	Design	N	Ages	Diagnostic criteria	Prevalence any disorder	Proportional morbidity†		
						Emotional	Behavioural	Hyper activity
Chandigarh (North India), 1991-1994 [428]	2-phase	963	4-11	ICD-10	5.0%	52%	56%	19%
Kerala (South India) 1992[427]	2-phase	1403	8-12	ICD-10	9.4%. (5.2% with impact criteria)	28%	70%	14%
Goa (South India), 2002-2003[426]	1-phase	2048	12-16	DSM–IV	1.8%	73%	27%	11%
Great Britain, 1999 & 2004 [B-CAMHS] [2-3]	1-phase	18 415	5-16	ICD-10	9.0%	47%	58%	17%

† Calculated as the number of children with each type of disorder as a proportion of all children with a common mental disorder. The figures sum to over 100% because of comorbidity.

Two studies report prevalences of common mental disorder substantially lower than B-CAMHS (1.8 and 5.0%, vs. 9.0%), and the third Kerala study likewise reported lower rates (5.2%) when an impairment criterion were used. The studies were less consistent regarding

[10] A fourth study has also been published, presenting results on a two-phase study of 1578 4-16 year olds in Bangalore conducted between 1995 and 2000 429. Srinath, S., et al., *Epidemiological study of child & adolescent psychiatric disorders in urban & rural areas of Bangalore, India.* Indian J Med Res, 2005. **122**(1): p. 67-79.. This study only presented the number of disorders identified among screen positive children, however, and gave no details regarding the screen negatives. It also did not allow calculation of proportional morbidity rates as it gave information only on the number of disorders of each type identified, and not on the total number of children affected. I was unable to obtain this information from the authors, and I was therefore unable to make use of the study.

proportional morbidity for different types of disorder.[11] One suggested relatively fewer behavioural disorders than in Britain [426], which is consistent with the evidence that the British Indian advantage may apply specifically to externalising disorders. The other two studies, however, showed a similar [428] or higher [427] proportionally morbidity from behavioural disorders.

4.1.5 Discussion

Limitations of this review

Before discussing my findings, it is worth highlighting some limitations of this review. Publication bias is particularly acute for routinely collected variables like ethnicity, as is the problem of relevant findings being 'hidden' in the main body of reports but not included in the abstract. Despite my multiple-pronged approach, I am therefore likely to have missed some studies, particularly those reporting null findings.

The heterogeneity of exposures and outcomes in this review made formal meta-analysis techniques impossible. I grouped studies by whether they reported statistically significant differences at the 5% level between minority groups and White British children as I felt this helped to clarify trends in the data. This method does, however, have several major limitations, including giving inadequate weight to studies reporting large and highly significant effects, and giving too much weight to underpowered studies reporting no effect. In addition, to avoid favouring studies including multiple testing on the same subjects, I presented each study only once, using combined categories such as 'fewer problems/no difference' where necessary. This does, however, give insufficient weight to studies showing consistent findings across multiple informants or by multiple measures. Like most meta-analyses, I also synthesised evidence without regard to variation in study quality.

A further, major drawback of the method I use is that it reinforces the idea that White/White British children represent an invariant, normative benchmark. This obscures important potential differences within White children both by ethnicity (e.g. migrants from

[11] The two-phase design of 2/3 studies meant I could not calculate prevalence estimates for different disorder types because the papers do not specify how many of each disorder type came from screen positive children and how many from screen negatives.

different European countries) and by other characteristics such as geographic region or SEP. Nevertheless, I hope to have reduced the problem of geographic and socio-economic variation somewhat by making all ethnic comparisons *within* studies or with comparable general population samples.

Key findings for common mental health problems

For common health problems, population-based studies suggest that Black African and Indian children may enjoy better mental health than White British children, while the mental health of Mixed race, Black Caribbean, Pakistani and Bangladeshi children is similar. For other minority groups there is insufficient evidence to draw conclusions. Evidence for a mental health advantage is particularly convincing in Indian children, given the comparatively large number of studies (N=12, including 10 non-B-CAMHS studies).

Within the common mental health problems, Indian children consistently displayed relatively more emotional and/or fewer behavioural problems, while the converse appeared true of Black Caribbean and Mixed White/Black Caribbean children.

The causes of these of inter-ethnic similarities and differences have been little investigated and remain largely unexplained. Only eight population-based studies of common mental health problems examined possible mediating or confounding factors. These include the B-CAMHS surveys [2-3] and two other large, recent studies, RELACHS [157] and DASH [413]. All four adjusted for various factors including individual child characteristics, child's identity and degree of assimilation, family structure, family social support, family activities and family coherence, various indicators of SEP and area deprivation. In three studies this had little effect on observed advantages and unmasked a relative advantage in other minority groups (namely 'Black' [157], Black Caribbean [413] and, in B-CAMHS04, Pakistani [3]). In the fourth study, B-CAMHS99, Indians were combined into a broader 'Asian' group in multivariable analyses [417], meaning the effect of adjustment on the univariable Indian advantage could not be assessed. In the case of the RELACHS study, further detailed analysis of a Bangladeshi advantage suggested that in girls the advantage was confined to those wearing more traditional clothing [430]. This finding is intriguing, suggesting the possible importance of cultural identity for mental health, but the precise mechanism is unknown. Instances of disadvantage in minority ethnic groups have also been little investigated, but may in some cases be explained by SEP [372], social support [387] and migration-related factors [374, 388].

Priorities for future research

This review reveals heterogeneity in the mental health of children from different ethnic groups, including within the groups 'Black' and 'South Asian'. It thereby highlights the importance of defining, reporting and analysing ethnicity in at least as much detail as the 2001 UK census [368].

This review also underlines the need for more sophisticated mental health evaluation in cross-cultural comparisons. Twenty-two of the 31 population-based studies in this review relied exclusively on brief mental health questionnaires, including 10/11 of the non-B-CAMHS population-based studies of Indians. As reviewed in Chapter 2, this may yield misleading findings if there are ethnic differences in the nature of mental health illnesses or in the perception and reporting of symptoms. Very few studies attempted to address this possibility. A more rigorous and more systematic approach to addressing the challenges of cross-cultural research is therefore needed. This should include using detailed interview-based measures in addition to questionnaires; examining the internal consistency of questionnaire subscales; comparing inter-informant agreement; and including a qualitative component to research projects.

As argued in Section 3.1.1 Chapter 3 (p.65), ethnicity is multi-faced construct which combines biological elements, ethnic self-identification, and broader social and structural factors [249]. Precisely for this reason, any observation of inter-ethnic differences should be a starting point for further hypothesis-driven investigations of causal mechanisms. Disappointingly few studies rose to this central challenge. For example, fewer than half (22/49) the studies in this review presented any data on SEP and only seven adjusted for SEP (despite differences being seen in 17/22 cases). In total, only 12 studies examined mediating or confounding factors which might explain ethnic differences. Most of these analyses suffered important limitations such as using broad ethnic groups like 'South Asian' (seven studies, including B-CAMHS99) or only presenting models adjusting for multiple factors simultaneously and so making it impossible to disentangle their individual contributions (four studies, including B-CAMHS99 and B-CAMHS04). As such, one of this review's most striking findings is that although a moderately large literature indicates ethnic variations in mental health, including an Indian mental health advantage, the causes of these variations remain largely unexplained.

4.1.6 Summary and conclusions

In summary, the prevalence of common mental health problems in the main minority ethnic groups in Britain seems to be similar to or, in some minorities, lower than that of White British children. This lack of a *dis*advantage is certainly reassuring, although parity with White British children still corresponds to a high burden of problems (with a mental disorder prevalence of almost 10%).

The evidence of a mental health advantage is particularly consistent for Indian children. This suggests that the observed Indian advantage which motivated this PhD is likely to reflect a genuine phenomenon rather than a chance finding. Furthermore, Indian children appear to be particularly advantaged for externalising problems – a finding which may explain the lack of any notable Indian advantage for adult common mental disorders (i.e. emotional disorders). It is also intriguing that recent population-based epidemiological surveys in India report much lower rates of common child mental disorders. Interpreting this finding is difficult given the many differences between India and Britain and between the three Indian study settings. Nevertheless, these findings raise the possibility that some aspect of Indian cultural or social life may be protective for child mental health and may be shared between Indian families in Indian and in Britain.

Yet unfortunately this review also highlights that there is little existing evidence to assess this hypothesis. The causes of the British Indian advantage, and of the other apparent inter-ethnic differences, remain unclear. Most existing research – including from B-CAMHS – contains very little investigation of the causal mechanisms underlying observed differences. Most studies also fail to demonstrate the validity and comparability of their mental health measures across ethnic groups, thereby undermining confidence in their findings. Addressing more fully these central challenges of cross-cultural psychiatric research is therefore crucial for understanding how and why mental health varies across ethnic groups. Doing this for the Indian advantage is the objective of this PhD, and the subsequent Chapters describe how I do this through secondary analysis of the B-CAMHS surveys.

Chapter 5 The British Child and Adolescent Mental Health Surveys 1999 and 2004

This Chapter introduces the British Child and Adolescent Mental Health Surveys (B-CAMHS) of 1999 and 2004 which provide the data analysed in this PhD. Section 5.1 and 5.2 describe the purpose and survey methods of B-CAMHS99 and B-CAMHS04. Section 5.3 evaluates the B-CAMHS mental health measures, namely a multi-informant semi-structured interview (the DAWBA) and a dimensional questionnaire (the SDQ). Finally, Section 5.4 then summarises the main findings from B-CAMHS to date.

5.1 Purpose of the B-CAMHS surveys

5.1.1 Historical context: previous child mental health surveys in Britain

The first child mental health studies using high-quality epidemiological methods were the Isle of Wight studies, as conducted by Michael Rutter and colleagues in the mid-1960s [68].[12] These were, and remain, very influential. This is not only because of their substantive findings, but because they pioneered the use of many core methodologies in child psychiatric epidemiology. These included using a defined, population-based sampling frame; employing questionnaires and standardised interviews of known reliability and validity; collecting data from multiple informants (parents, teachers and children); and using clinician input when making diagnoses.

The Isle of Wight studies also provide an early example of a multiphase-multimethod survey design in child mental health. These typically employ a two-phase design, with all children being administered brief screening questionnaires in the first phase and a subset of

[12] In fact, Rutter had previously collaborated in the Aberdeen Child Development Study, a population-based cohort of over 12 000 children initiated in 1962-64 431. Leon, D.A., et al., *Cohort profile: the Aberdeen children of the 1950s study.* Int J Epidemiol, 2006. **35**(3): p. 549-52.. It was in this survey that he validated the parent and teacher Rutter questionnaires but although responses to individual items have been linked to later outcomes the data has never been analysed using the full Rutter questionnaires 65. Henderson, M., M. Hotopf, and D.A. Leon, *Childhood temperament and long-term sickness absence in adult life.* British Journal of Psychiatry, 2009. **194**: p. 220-223..

children being administered more intensive interviews in the second. These intensive interviews are usually applied to all children who screen positive and a randomly selected sub-sample of those who screen negative. In the Isle of Wight studies, children were screened using the Rutter questionnaires for parents and teachers. Mothers of all children screening positive and one in 12 of those screening negative were then interviewed in detail by a child psychiatrist, using a standardised procedure and blinded to whether the child had screened positive or negative. This oversampling of children at higher risk of problems offers potential gains in efficiency, by permitting reasonably reliable prevalence estimates without having to administer in-depth interviews to each child. Multiphase-multimethod designs may therefore permit larger sample sizes and/or more detailed assessment methods (e.g. clinician interviews) than one-phase designs.

Many subsequent studies have employed the Isle of Wight methodologies. This is shown in Table 5.1 which presents information on British population-based studies with a sample size of over 1000 children. I located these studies through a review by Verhulst and Koot [4] which I updated to include more recent surveys. There were nine studies other than the Isle of Wight and B-CAMHS surveys. Four were nationally representative, the remainder using total local population samples or representative local samples. Six studies used multiple informants, usually parents and teachers but in two cases children as well. Three studies used two-phase designs and had clinician input in making diagnoses. The remaining six used one-phase designs and only used questionnaire measures.

Most studies estimated prevalences of 10-15% for mental disorder or ‘high’ questionnaire scores, in line with surveys from other parts of the world [4].

Table 5.1: Methodological features of large population-based studies of child mental health in the UK

Author	Study date	Location	N†	Age	Sampling strategy	Sampling frame	Design	Informant(s)	Type(s) of measure	Validated measure ††	Clinician input?	Estimated prevalence of mental health problems/disorders
Shepherd *et al.* [432]	1961	Buckinghamshire	6411	5 to 15	Locally representative sample	Schools	One-phase	Parent; Teacher	Questionnaire	No	No	Not assessed
Rutter *et al.* [68, 433]	1964-5	Isle of Wight	1279	10	Total local population	Schools	Two-phase	Parent; Teacher; Child	Questionnaire; Interview	Yes	Yes	12.0% clinical disorder
Leslie [434]	1968	Blackburn	1198	13 to 14	Total local population	Schools	Two-phase	Parent; Teacher; Child	Questionnaire; Interview	Yes	Yes	20.8% clinical disorder boys, 13.6% girls
Rutter *et al.* [433]	1970	Inner London	1689	10	Total local population	Schools	Two-phase	Parent; Teacher	Questionnaire; Interview	Yes	Yes	25.4% clinical disorders
Mensah *et al.*[435] [NCDS]	1965, 1969 and 1974	England, Scotland and Wales	11036	7, 11 and 16	Nationally representative sample	Whole population birth cohort	One-phase	Parent; Teacher	Questionnaire	No†††	No	5.0% with high scores at every sweep on parent questionnaire; 15.0% with high scores at one or two sweeps.
Davie *et al.* [436] [NCDS]	1965	England, Scotland and Wales	15425	7	Nationally representative sample	Whole population birth cohort	One-phase	Parent; Teacher	Questionnaire	Yes	No	14% high scores on teacher questionnaire
Buchanan *et al.* [437] [NCDS]	1974	England, Scotland and Wales	8441	16	Nationally representative sample	Whole population birth cohort	One-phase	Parent; Teacher	Questionnaire	Yes	No	7.2% high score on parent questionnaire
Mensah *et al.*[435] [BCS]	1975, 1980 and 1986	England, Scotland and Wales	10653	5, 10 and 16	Nationally representative sample	Whole population birth cohort	One-phase	Parent; Teacher	Questionnaire	No†††	No	3.4% high score at every sweep by the parent questionnaire; 12.6% high score at one or two sweeps.
Collishaw *et al.* [7] [BCS]	1986	England, Scotland and Wales	7293	16	Nationally representative sample	Whole population birth cohort	One-phase	Parent; Teacher	Questionnaire	Yes	No	10.5% high score on parent questionnaire for emotional problems; 10.4% for behavioural; 7.1% for hyperactivity.

Author	Study date	Location	N†	Age	Sampling strategy	Sampling frame	Design	Informant(s)	Type(s) of measure	Validated measure ††	Clinician input?	Estimated prevalence of mental health problems/disorders
Stallard [438]	1990	Bath	1170	3	Total local population	Heath clinics	One-phase	Parent	Questionnaire	Yes	No	10% high score; 16% parents report a lot of concern
McMunn *et al.* [439] [HSE 1995-7]	1995, 96 & 97	England	5705	4 to 15	Nationally representative sample	Postcode address file	One-phase	Parent	Questionnaire	Yes	No	12% high scores boys, 8% girls
Meltzer *et al.* [2] [B-CAMHS99]	1999	England, Scotland and Wales	10 438	5 to 15	Nationally representative sample	Child benefit registry	One-phase	Parent; Teacher; Child	Questionnaire; Interview	Yes	Yes	11.4% clinical disorder boys; 7.6% girls
Sproston *et al.* [440] [HSE 2001/2]	2001 and 2002	England	5882	4 to 15	Nationally representative sample	Postcode address file	One-phase	Parent	Questionnaire	Yes	No	12% high scores boys, 8% girls
Green *et al.* [3] [B-CAMHS04]	2004	England, Scotland and Wales	7 974	5 to 16	Nationally representative sample	Child benefit registry	One-phase	Parent; Teacher; Child	Questionnaire; Interview	Yes	Yes	11.4% clinical disorder boys; 7.8% girls

NCDS=National Child Development Study (birth cohort of 1958); BCS=British Cohort Study (birth cohort of 1970); HSE=Health Survey for England
† Sample sizes correspond to children for whom mental health data is available or, if this is unclear, for the total sample. †† A validated measure was defined in the same way as in my systematic review (Chapter 4, p.94). ††† The two studies by Mensah *et al*. used only 11 items from the Rutter questionnaire, these having been chosen on the basis of exploratory factor analyses but without validation [441]. Most other analyses from NCDS and BCS only report individual Rutter items [e.g. 442] or report mean scores stratified by various risk factors [e.g. 443]. Only for the NCDS sweep at age 7 and the NCDS and BCS sweeps at age 16 could I find published analyses reporting the proportion of high-scores.

5.1.2 Aims of the B-CAMHS surveys

The B-CAMHS surveys were thus predated by several large, population-based studies. Some included children of all ages, some were nationally representative, and some had clinical assessments of mental disorder. No previous survey, however, included all these elements.

The rationale for B-CAMHS was therefore to bring these strengths together to provide the most comprehensive child mental health surveys ever conducted in the UK [2-3]. The key aims were:

1. To estimate the prevalence of emotional, behavioural and hyperactivity disorders of children living in private households. B-CAMHS04 also included expanded sections on less common disorders, in order to assess their prevalence more accurately.
2. To determine the impact and burden of child mental disorders.
3. To examine the use of mental health services among children with mental disorders.

In addition, the B-CAMHS surveys aimed to investigate the child, family, school and area correlates of child mental disorders. The inclusion of three-year follow-up also allowed assessment of disorder persistence, prognosis and onset.

The B-CAMHS surveys were initiated and funded by the Department of Health. They were conducted by the Office for National Statistics (ONS) in collaboration with the Institute of Psychiatry, London.

5.2 Sampling and survey procedures in B-CAMHS

5.2.1 The B-CAMHS baseline surveys

Sampling frame

The B-CAMHS surveys were representative, population-based surveys of children aged 5-15 (B-CAMHS99) or 5-16 (B-CAMHS04) in England, Scotland and Wales [2-3]. Both used the Child Benefit Centre (CBC) register of the Department of Work and Pensions as a sampling frame. This had the advantage of allowing children to be selected with equal probabilities, avoiding the weights required when sampling one child per household using Postcode Address Files. As compared to school-based sampling, the CBC register also had the advantage of facilitating the collection of parental consent for approaching teachers.

The CBC register covers nearly 99% of children in Great Britain [444]. Of these, the B-CAMHS sampling frame excluded children without postcodes (10% in 1999 and 2% in 2004). The B-CAMHS team report that the CBC had no evidence that records with postcodes differed from those without [2]. The B-CAMHS sampling frame also excluded those for whom the CBC was invoking administrative action; the size of this group is not given, but was described as a "small minority" by one of the lead ONS investigators (Rebecca Gatward – personal communication). The B-CAMHS team argues that the administrative nature of this action means excluding these child "should not bias the sample in any way" [2, p.22]. This seems unduly optimistic, however, as reasons for action include the sensitive nature of the case, a change of address or the death of the child. It is highly plausible that children with sensitive cases constitute a non-random sample in terms of their mental health. This may also be true of children who have recently changed address. Unfortunately, it is unknown how far this may compromise the representativeness of B-CAMHS.

Sampling design

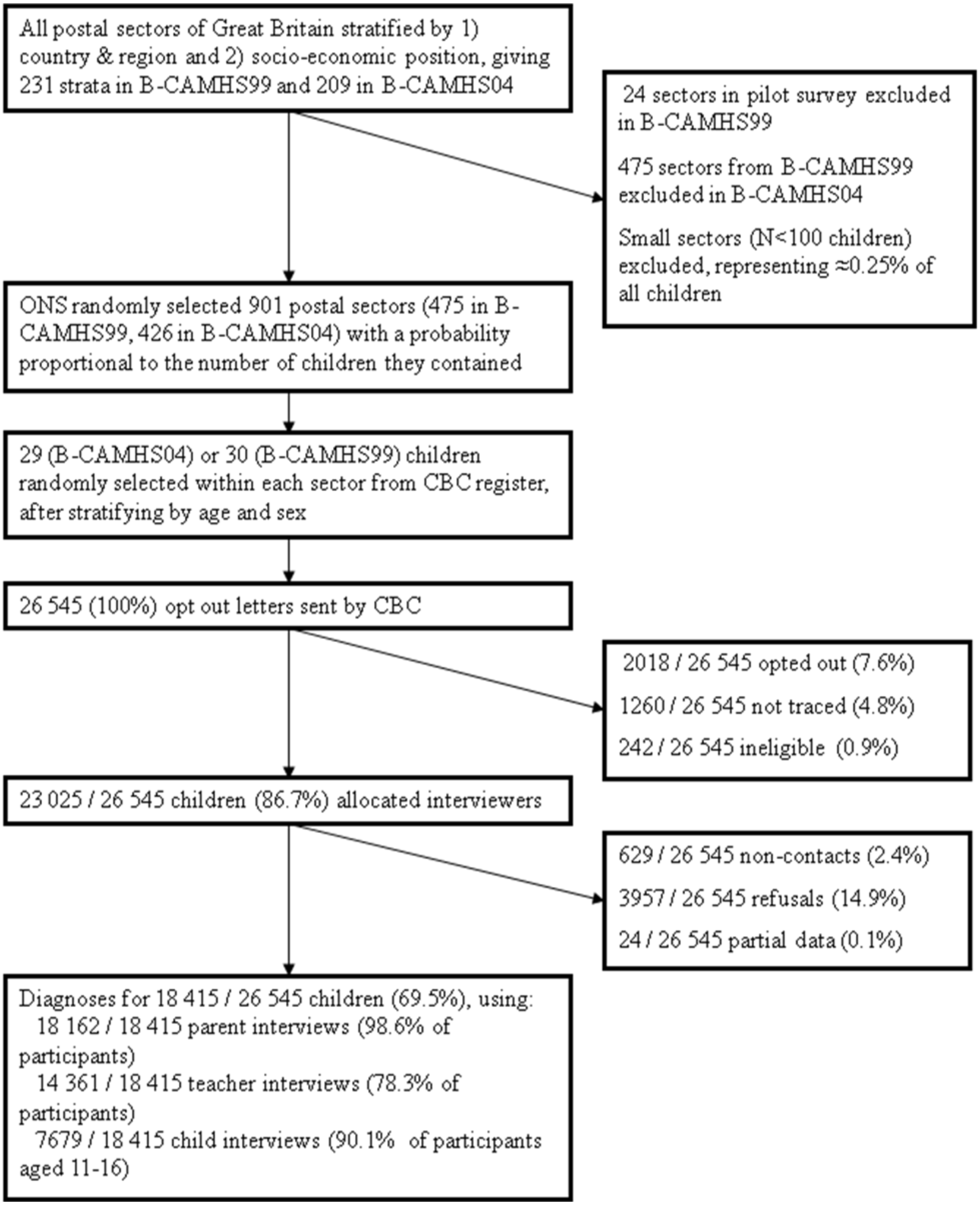
Figure 5.1 summarises the recruitment of children into B-CAMHS. Both surveys employed a complex sampling design using stratification, clustered sampling and weighting. I present a general discussion of the principles underlying these techniques in Section 13.3 Appendix 1, and in Section 6.1 Chapter 6 I discuss how I deal with this complex design in my analyses.

The primary sampling units in both B-CAMHS surveys were postal sectors. First, postal sectors in Britain were stratified by geographical region and then by socio-economic group (for details see [2 (Appendix A), 3 (Appendix A)]). This generated a total of 231 strata in B-CAMHS99 and 208 in B-CAMHS04. Postal sectors containing fewer than 100 children were then excluded, corresponding to approximately 0.25% of children on the CBC register. The remaining postal sectors were sampled at random with a probability proportionate to the number of children on the CBC register aged 5-15 (B-CAMHS99)/5-16 (B-CAMHS04). In B-CAMHS99 strata from Wales and the Scottish Highlands and Islands were slightly oversampled, while in B-CAMHS04 only a half-sample was taken in Wales for financial reasons. Postal sectors included in B-CAMHS99 were excluded from the B-CAMHS04 sampling frame.

Two postal sectors each were selected from 218/231 strata in B-CAMHS99 and 198/208 in B-CAMHS04. Three postal sectors each were selected from the remaining 13 strata in B-CAMHS99 and 10 strata in B-CAMHS04. This resulted in a sample of 901 postal sectors (475 in B-CAMHS99 and 426 in B-CAMHS04), out of a total of 8265 in Great Britain. At the request of ONS, the CBC stratified within each postal sector by age and sex and generated a random sample of 30 children per sector in B-CAMHS99 or 29 children per sector in B-CAMHS04. In B-CAMHS04, 5/426 postcodes contained fewer than the required numbers of families with children in the target age range, leading to an overall shortfall of 60 children. In addition, one child too few was selected in one sector in B-CAMHS99, while one child too many was selected in two sectors in B-CAMHS04. The total number of children selected was therefore (30*475)+(29*426)-60-1+2=26 545.

In combination with sampling postal sectors with a probability proportionate to their size, the selection of a fixed number of children per postal sector means that each child in the CBC sampling frame had an equal chance of being selected. Probability weights were, however, subsequently calculated to adjust for over- or under-sampling of strata in different countries in Britain, and for differential non-response by region, age and sex.

Figure 5.1: Overview of recruitment into the B-CAMHS surveys



Recruitment of children into the surveys

The CBC sent the parents/guardians of the 26 545 selected children a letter informing them about the survey and giving them an opportunity to opt out. ONS then received the names and addresses of parents who did not opt out and approached them to participate in the study. With parental consent, ONS also approached the child themselves if aged over 11 and one of their teachers (see Box 5.1). Interviews with parents and children were conducted face-to-face, while teachers were sent printed interviews to complete. All eligible children were administered the full range of questionnaire and interview measures regardless of their mental health status (i.e. a one-phase design). This process is presented in the bottom half of Figure 5.1, together with the reasons for non-participation.

Overall it was possible to make a DAWBA diagnosis for 73.2% of the original sampling frame in B-CAMHS99 (10438/14249) and 64.9% (7997/12296) in B-CAMHS04. Table 5.2 summarises the age, sex and ethnic group of these children. In Section 6.3 Chapter 6, I present detailed analyses of whether these and other characteristics predicted non-participation among parents, teachers and children.

Box 5.1: Participants interviewed in the B-CAMHS surveys

- **Parents**: For each child selected from the CBC register, the research team approached one parent or guardian to act as a primary informant. 93% of these parents or guardians were the child's mother. Throughout this PhD I use 'parent' to refer to the interviewed parent and 'parent's partner' for the non-interviewed parent. Information was collected from parents by face-to-face interview.
- **Children:** Participating parents were asked for permission to approach children aged 11 or over and to invite them to take part. If the child consented, information was collected from them by face-to-face interview. In addition, in B-CAMHS99 children of all ages were (with consent) administered a reading and vocabulary assessment.
- **Teachers**: Participating parents were asked for permission to approach the teacher whom the parent considered 'knew the child best'. Information was then collected from teachers by a postal interview.

Table 5.2: Characteristics of participating children by age, sex and ethnicity

		B-CAMHS99 (N=10438)	B-CAMHS04 (N=7977)	Combined B-CAMHS surveys (N=18415)
Age†	**5-7 yrs**	2964 (28.4%)	1920 (24.1%)	4884 (26.5%)
	8-10 yrs	2949 (28.3%)	2005 (25.1%)	4954 (26.9%)
	11-13 yrs	2790 (26.7%)	2130 (26.7%)	4920 (26.7%)
	14-15/16	1735 (16.6%)	1922 (24.1%)	3657 (19.9%)
Gender	**Male**	5212 (49.9%)	4111 (51.5%)	9323 (50.6%)
	Female	5226 (50.1%)	3866 (48.5%)	9092 (49.4%)
Ethnicity (grouped) ††	**White**	9529 (91.4%)	6920 (86.7%)	16449 (89.4%)
	Black	247 (2.4%)	198 (2.5%)	445 (2.4%)
	Indian	215 (2.1%)	199 (2.5%)	414 (2.2%)
	Pakistani & Bangladeshi	189 (1.8%)	307 (3.8%)	496 (2.7%)
	Other groups	251 (2.4%)	349 (4.4%)	600 (3.3%)

†Children sampled to age 15 in B-CAMHS99 and 16 in B-CAMHS04. †† Ethnicity data missing on 11 children.

5.2.2 Overview of data collected in B-CAMHS

Mental health outcomes

The B-CAMHS surveys collected mental health data using the semi-structured interview the Development and Well-Being Assessment (DAWBA) and the brief Strengths and Difficulties Questionnaire (SDQ). A detailed description of these measures and their psychometric properties is provided in the next section of this Chapter.

Potential correlates of child mental health problems

The B-CAMHS surveys collected a variety of information about child, family, school and area characteristics. Ethnicity is the main explanatory variable of interest in this PhD, and Section 6.2.2 Chapter 6 describes and evaluates how ethnicity was measured and how I use it in this PhD. The other potential explanatory variables collected are summarised below, and described in detail in Section 9.2, Chapter 9

- **Child characteristics**: Age and sex; General health; Specific health complaints and physical disorders; Stressful life events; Substance use (smoking, alcohol, drugs); Learning difficulties and dyslexia; Formal assessments of academic ability; Parent assessment of academic ability; Teacher assessment of academic abilities; Parent's use of rewards and punishments; Parent's opinion of friends; Social aptitudes scale; Social support; Number of close relatives; Helping relatives out.

- **Family stress**: Parent's mental health; Family functioning; Stressful life events affecting the family.
- **Family composition**: Family type (two-parent vs. lone parent vs. step family); Parent marital status (married vs. cohabiting); Three generation family; Number of resident siblings; Mother's age at child's birth.
- **Family socio-economic position**: Parent's education; Household income; Rented housing tenure; Occupational social class; Mother's economic activity; Father's economic activity.
- **School characteristics**: Ford score (a school-level predictor of the prevalence of child mental health problems in schools).
- **Area characteristics**: Country and region; ACORN geodemographic classification; Carstairs measure of deprivation; Index of Multiple Deprivation measure of area deprivation; Area ethnic density.

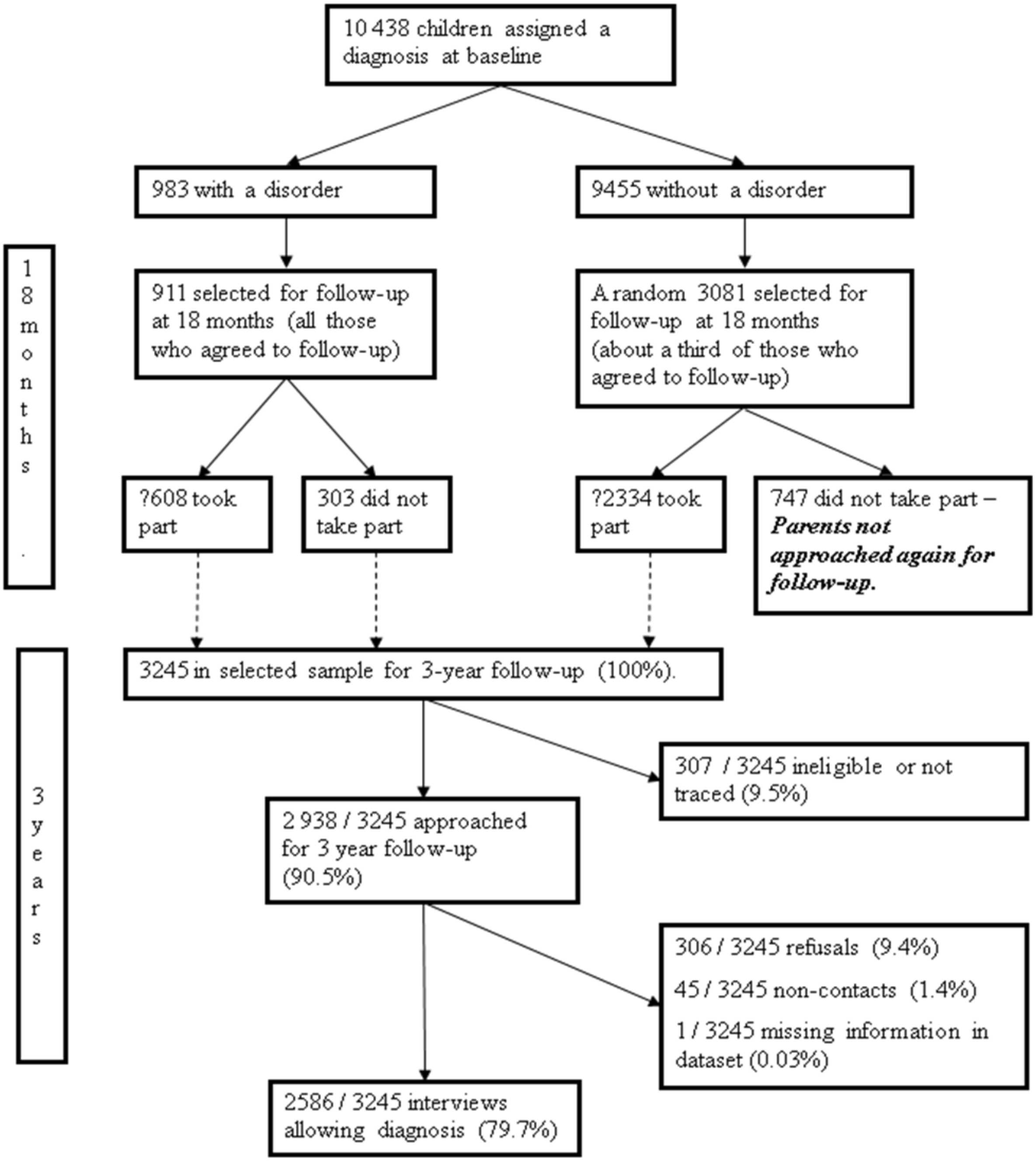
5.2.3 The B-CAMHS follow-up surveys

Both B-CAMHS surveys collected interim follow-up data at various points in the three years following the original study. SDQ scores were collected at six and 18 months in B-CAMHS99, and at six, 12 and 24 months in B-CAMHS04. SDQ's were collected by postal surveys or, for those who did not respond, over the telephone. These interim follow-ups varied in whether they recruited teachers or only parents, and in whether they approached all parents or just a random subsample (for details, see [48, 445]). I do not use this interim SDQ data in this PhD.

In both surveys, a full re-assessment of mental health and risk factors occurred after three years by face-to-face interview, with information again collected from parents, teachers and children aged over 11 [48, 445]. In B-CAMHS99, a rather complicated process was used to select children for follow-up. Children were approached 1) if the child had been diagnosed with a disorder at baseline or 2) if the child had not been diagnosed at baseline but at least one informant had returned a questionnaire at 18 months (Figure 5.2). Furthermore, 18 month follow-up had been attempted only for a random third of those without a disorder at baseline. Three-year follow-up was therefore attempted for all children with a disorder at baseline, but only for a randomly selected third of disorder-free children who had also returned questionnaires at 18 months. By contrast, in B-CAMHS04 all children were approached for three-year follow-up (Figure 5.3).

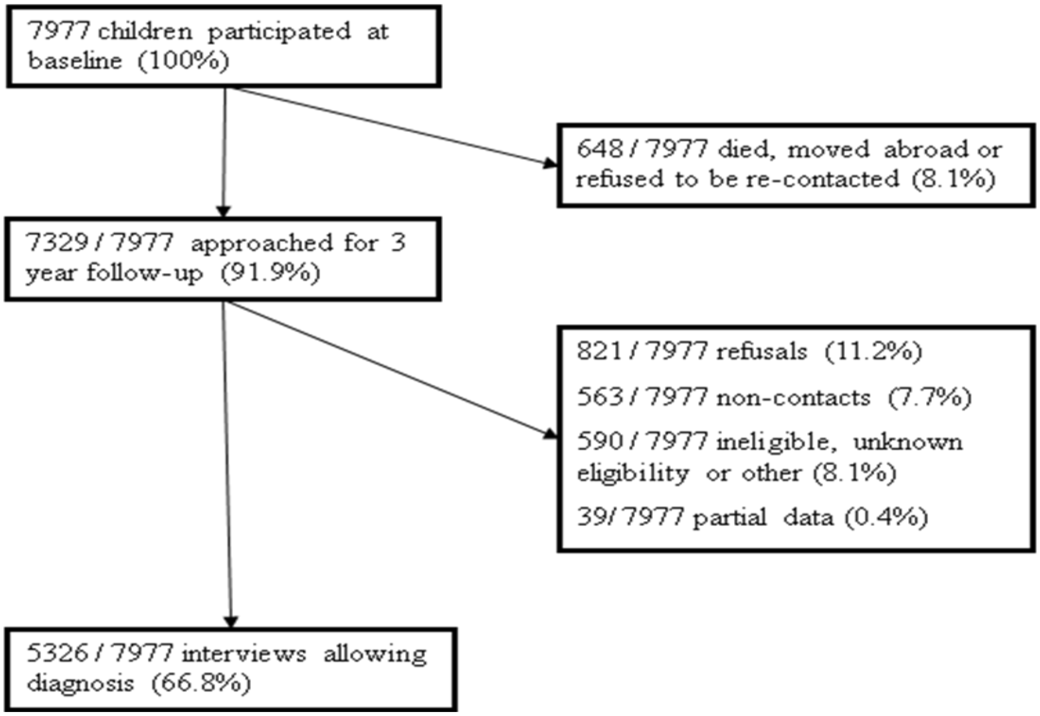
Overall, three-year follow-up information was collected on 2586 children from B-CAMHS99 (79.7% of those approached) and 5326 children from B-CAMHS04 (66.8% of those approached). In both B-CAMHS surveys, 95% of interviews at three years were with the same parent who had been interviewed at baseline, with mothers making up 94% of all parents interviewed.

Figure 5.2: Recruitment of participants in B-CAMHS99 to three-year follow-up



The official B-CAMHS publication contains minor numerical inconsistencies regarding the 18 month questionnaires, which ONS could not clarify. Figures preceded by a question mark are therefore inferred.

Figure 5.3: Recruitment of participants into follow-up for B-CAMHS04, at 6, 12 and 24 months, and three years.



5.2.4 Strengths and limitations of the B-CAMHS survey design

<u>Strengths</u>

Only a small and non-representative subset of children with mental health problems access mental health services [11, 68, 446]. Population-based sampling is therefore essential for achieving representative samples and accurate prevalence estimates. Moreover, the potential for differential referral patterns in different ethnic groups [404, 409, 447] means that population-based sampling may be particularly important for ethnic comparisons [243, 360]. That B-CAMHS provides such data is therefore a central strength in addressing the aims of this PhD.

As summarised previously in Table 5.1, B-CAMHS is not the first large, population-based study to collect data on child mental health. The B-CAMHS surveys do, however, have several major strengths over the previous population-based surveys. Firstly, they are much larger: only one previous survey had a sample size of over 5000. Moreover, B-CAMHS covered a wider age range (5 to 15/16), took a nationally representative sample, and used clinical assessment in assigning disorders. All four previous nationally-representative samples in Britain used only brief questionnaires, while the three previous surveys using clinical assessment were based only in particular areas.

The three previous surveys using clinical assessments also all used a multiphase-multimethod design, rather than the one-phase design of B-CAMHS. Multiphase-multimethod designs are widely used in child psychiatric epidemiology [4, 239] due to their potential to enhance efficiency. Yet this increased efficiency carries some costs for precision and validity. Prevalence estimates in multiphase-multimethod designs involve extrapolation from a subsample of screen-negative children, thereby increasing sampling error. Multiphase designs also contain additional potential for bias because there is usually some internal loss of children between screening and interview. By contrast, there are very few incomplete interviews in B-CAMHS (for example, 1.4% among parents). The availability of complete mental health information on all children also permits a detailed examination of emotions and behaviours across the full range of mental health functioning. This is particularly relevant to the focus of this PhD upon protective factors.

A final strength of the B-CAMHS survey design is the inclusion of a three-year follow-up. As described above, whereas B-CAMHS04 attempted complete follow-up of all children, B-CAMHS99 approached under a third of children without disorder at baseline. Nevertheless, the large baseline samples mean that large follow-up samples are available from both surveys (2586 from B-CAMHS99 and 5326 from B-CAMHS04). Moreover, the follow-up data collected was unusually comprehensive, including mental health assessments by parents, teachers and children, and also a re-assessment of many of the potential explanatory variables of child mental health problems. This allows some exploration of the causal directions between child mental health and its correlates, an issue to which I return in Section 11.1.2, Chapter 11.

Limitations

The B-CAMHS surveys provide the most comprehensive assessment of child mental health ever conducted in Britain, but do also have several important limitations. One is the scope for selection bias resulting from participation rates of only about 70%; in Section 6.3 Chapter 6 I present analyses which examine this issue.

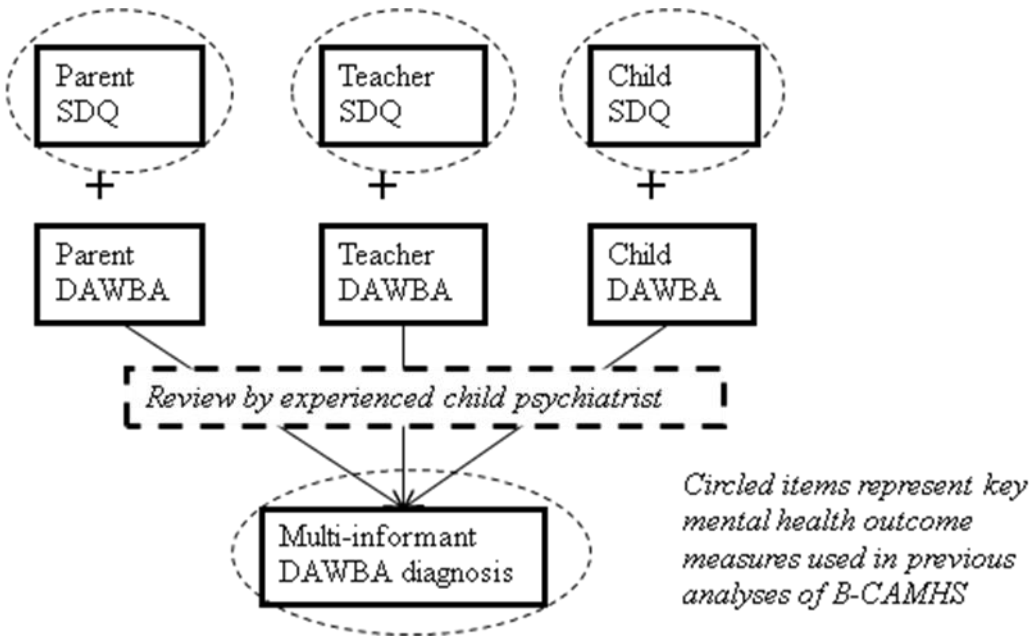
From the perspective of this PhD, another limitation is the absence of oversampling of minority ethnic groups. This means that while the total B-CAMHS sample size is large, the number of individuals in many ethnic groups is small (N=419 for Indians). I return to the implications of this in subsequent Chapters.

5.3 Mental health measures used in B-CAMHS

5.3.1 Overview of data collection on mental health

All parents, teachers and children participating in B-CAMHS were administered two mental health measures. First, they were administered the Strengths and Difficulties Questionnaire (SDQ), a dimensional questionnaire measure of child mental health. Next they were administered the DAWBA interview, a diagnostic interview containing both structured sections of closed questions and open-ended transcripts. All informants were administered the DAWBA regardless of their SDQ score (although as described below, SDQ scores did form one part of the skip rules used in administering some sections of the DAWBA). Experienced child psychiatrists then used the closed and open-ended responses to the DAWBA from all available informants to make a single, global decision about whether a child merited a clinical diagnosis. Multiple mental health measures are therefore available for each child, as illustrated in Figure 5.4.

Figure 5.4: Schematic illustration of mental health assessment in the B-CAMHS surveys



In this section, I describe the SDQ and DAWBA in detail, and place them in the context of other questionnaire and interview measures. I then summarise and evaluate the evidence on their psychometric properties, with a particular focus upon evidence from the UK general population. Finally I introduce and describe a preliminary evaluation of the DAWBA bands. These are measures of symptoms reported in the parent, teacher and child DAWBA interviews. The DAWBA bands therefore bridge the gap between the informant-specific SDQ scores and the multi-informant DAWBA diagnosis, these latter two being the outcomes used in all previous B-CAMHS analyses (Figure 5.4).

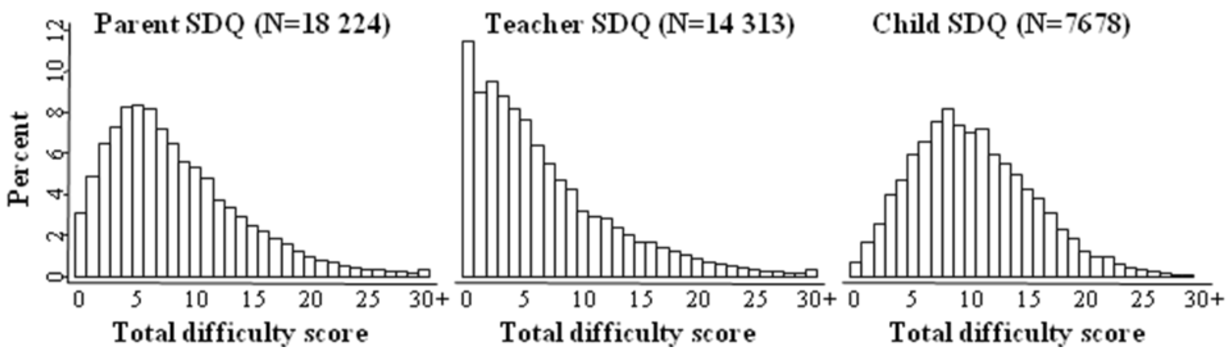
5.3.2 The Strengths and Difficulties Questionnaire (SDQ)

Description of measure

The Strengths and Difficulties Questionnaire (SDQ) is a 25-item questionnaire with supplementary questions on distress and impairment. It can be administered to parents and teachers of children aged 4 to 17, and to children aged 11 or over.

The SDQ was designed to include five subscales relating to emotional problems, behavioural problems, hyperactivity, peer problems and prosocial behaviour [194, 448]. Each subscale is comprised of five items with ‘Not true’ (scored 0), ‘Somewhat true’ (scored 1) and ‘Certainly true’ (scored 2) as response options, giving a range of 0-10. Five of the 20 items on the emotional, peer, behavioural and hyperactivity subscales are ‘strenghts’ and are reverse scored (the five prosocial items are also all positively worded). Responses to the 20 items from the emotional problems, peer problems, behavioural problems and hyperactivity subscales are added to give a total difficulties score (TDS), with a range of 0-40. An annotated copy of the parent version of the SDQ is included in Section 14.2.1 Appendix 2, together with a summary of the constituent SDQ items of the five subscales and TDS. All scales (and particularly those from teachers) are usually positively skewed to some extent, as illustrated in Figure 5.5 using the B-CAMHS data.

Figure 5.5: Distribution of the parent, teacher and child TDS in B-CAMHS



In addition to these 25 symptoms, the SDQ contains supplementary questions on impact. The parent and child SDQs contain one item on distress to the child, and four items on interference in home life, friendships, classroom learning and leisure activities. The response options are ‘Not at all’/’Only a little’ (both scored 0), ‘Quite a Lot’ (scored 1) and ‘A great deal’ (scored 2), giving an impact score of 0-10. The teacher SDQ contains one item on distress to the child and two on interference with peer relationships and classroom learning, giving an impact score of 0-6. There is also a single four-point item about burden to others for all informants.

Comparison with other child psychiatric brief questionnaires

There exist several other brief questionnaires which provide dimensional measures of child mental health. Among the longest-established are the Rutter scales for parents and teachers [68], which have proved reliable, valid and useful in a wide range of settings [193]. Also widely used are the Achenbach System of Empirically Based Assessment for parents, teachers and children [ASEBA: 187, 188-191], which I introduced in Chapter 2, p.48.

The development of the SDQ started with an extended version of the Rutter questionnaire which included ‘strengths’ items [449]. The SDQ was then refined by using factor analyses to identify its five key subscales; by including items which explicitly relate to core diagnostic symptoms (e.g. concentration, restlessness and impulsiveness for hyperactivity); and by omitting behaviours like thumb-sucking which are no longer of interest to mental health professionals [194]. The result is that the SDQ has fewer items on emotional and behavioural problems than the Rutter; covers hyperactivity in a way which corresponds more closely to current diagnostic criteria; and has greater coverage of peer problems and prosocial behaviour. Unlike the Rutter, there also exists a child-report version of the SDQ.

The SDQ differs from the Rutter and the ASEBA in other important ways. The SDQ includes a substantial number of strengths items, as opposed to the exclusive focus of the Rutter and the ASEBA upon difficulties. The parent, teacher and child SDQs have identical items, facilitating comparisons across informants. By contrast, there are important differences in the content of the parent and teacher Rutter, and the parent, teacher and child ASEBA questionnaires. Finally, at 25 items the SDQ is slightly shorter than the Rutter (31 items in the parent version, 26 in the teacher version) and substantially shorter than the ASEBA (120 in the parent and teacher versions, 105 items on the child version). This may explain some evidence that the SDQ is more popular than the ASEBA with parents [450-451].

The reliability and validity of the SDQ total difficulties score (TDS) and subscales in Britain

Total difficulty score

There is substantial evidence that the SDQ’s total difficulty score (TDS) provides a reliable and valid measure of child mental health problems in Britain, with psychometric properties comparable or superior to other brief questionnaires. This evidence is summarised below and described in full in Table 14.8 of Section 14.2.2 Appendix 2.

Evidence of reliability of the TDS includes reasonable test-retest reliabilities (0.65-0.85) [452-453] and high Cronbach’s alphas (≥0.80 in B-CAMHS99) [448]. Agreement between parents, teachers and children was only moderate in B-CAMHS99 (0.33-0.48) [448] but nonetheless substantially better than the mean agreement reported in a meta-analysis of previous measures (0.20-0.27) [226]. Moreover, as discussed in Chapter 2 (p. 60), these comparatively low rates of agreement do not necessarily simply reflect poor measurement – they may also result from substantive differences in how a child behaves in different settings.

The TDS also shows good evidence of construct validity. Despite being briefer, the TDS is highly correlated (0.78-0.92) with the parent and teacher Rutter and parent ASEBA and functions at least as well at detecting high risk groups [194, 451]. 'High' scores show good sensitivity (≈80-90%) and specificity (≈80%) relative to clinical diagnosis [454-455]. Moreover, in the combined B-CAMHS surveys the TDS seems to represent a truly dimensional measure of mental health problems, with each one-point increase in parent, teacher and child TDS being associated with an increased prevalence of DAWBA diagnosis [456].[13] Further evidence to this effect comes from the fact that the same risk factors that predict change in TDS across the entire range, also predict it in children one standard deviation above and one standard deviation below the mean [457]. These risk factors are, moreover, the same as for the DAWBA and are in line with the previous literature [458].

SDQ subscales

There has been less evaluation of the reliability and validity of the five hypothesised SDQ subscales and the evidence is more mixed (see Table 14.9, Appendix 2). Cross-informant correlations are again not high but better than average for measures of child mental health, ranging from 0.25-0.48 for parent-teacher correlation, 0.30-0.44 for parent-child correlation and 0.21-0.32 to teacher-child correlation. The Cronbach alphas for some subscales are somewhat low (0.60-0.70), which may partly reflect the small number of items per scale [448]. The subscales showed evidence of good convergent and discriminant validity relative to DAWBA diagnoses in B-CAMHS99 [459]. The correlation with scores on the PACS investigator-based, semi-structured interview [460] are also least as high for the SDQ as for the ASEBA [451]. The subscales show the expected pattern of cross-scale correlation, with the behavioural and hyperactivity subscales being more highly correlated with each other than with the emotional subscale [448, 451]. Principal component analyses of B-CAMHS99 also broadly confirm the expected factor structure; in five-factor principal

[13] On this paper I was the lead author; see Appendix 3.

component analyses all 25 items loaded onto their expected scale in all three informants, with 72/75 of these loadings being >0.4 [448].

Chief limitations of existing psychometric evaluations of the SDQ

Yet while these principal component analyses are reassuring, the existence of a proposed factor structure means that confirmatory factor analyses (CFA) would be a more appropriate technique. This is because CFA provide a hypothesis driven and model-based framework which can formally test the relative fit of different models (see Appendix 1, Section 13.1.5 for a fuller discussion).

This gap in the evidence is particularly important because the international literature does not fully support the proposed five-factor structure. As in Britain, most international studies are limited to exploratory principal component analyses or exploratory factor analyses (EFA). Generally these support the proposed five-factor solution (for a recent review, see [42]). At least one EFA from Finland [461], however, and two CFA from the US [462] and Belgium [463] suggest a three-factor solution. These three factors comprise an ‘internalising’ scale of the emotional and peer items, an ‘externalising’ scale of the behavioural and hyperactivity items and the standard ‘prosocial’ subscale. Table 14.7 (Section 14.2.1 Appendix 2) summarises this alternative three-factor structure alongside the hypothesised five-factor structure.

Finally, a thorough CFA on 914 Australian parent, teacher and child SDQs found that the five-factor solution did not provide adequate fit to the data and that many items loaded onto multiple factors [464]. The authors note that many psychological scales fail to meet rigorous psychometric criteria, and that model fit was near-acceptable by some indices. Nevertheless, this study raises further grounds for caution regarding the SDQ factor structure, particularly since a CFA on 4167 Norwegian child SDQs also found a questionable model fit [465].

I therefore evaluate the SDQ’s internal structure further in Section 8.1 Chapter 8, conducting CFA and other analyses in order to inform my subsequent comparison of Indians and Whites.

Evaluation of the SDQ outside Britain

The satisfactory psychometric properties and clinical utility of the SDQ has been demonstrated in many high-, middle- and low-income settings including the US, Italy, Scandinavia, Brazil and the Yemen (for reviews, see [3 (Appendix D), 191, 200]). In India, the SDQ has not been validated but has been used with apparent success in Goa [426]. The SDQ has also been evaluated in other parts of the Indian subcontinent. Specifically, studies in Bangladesh [466-467] and Pakistan [468] have demonstrated that the SDQ differentiates clinic and community samples, strongly predicts clinical diagnosis, and can discriminate between different types of disorder within the clinic sample. Particularly noteworthy is the simultaneous evaluation of the multi-informant SDQ algorithm for probability of disorder in 101 children from a London clinic and 89 from a clinic in Dhaka, Bangladesh [454]. In both settings, the category of 'probable' diagnosis correctly identified most children (81-91%) who were independently assigned a clinical diagnosis, and the observed association was as strong in Dhaka as in London.

5.3.3 The DAWBA

Description of measure

The DAWBA interview

The Development and Well-Being Assessment (DAWBA) interview is a semi-structured interview administered to parents and teachers of children age 4-16, and to children over the age of 11 [459]. The DAWBA can either be administered by trained lay interviewers (used in B-CAMHS for parents and children) or else it can be self-completed (paper self-completion used in B-CAMHS for teachers). Child psychiatrists then use responses from across all informants to assign psychiatric diagnoses. Although it has since been used elsewhere, the DAWBA was initially developed specifically for B-CAMHS [2].

The DAWBA interview consists of a mixture of open and closed questions about child mental health symptoms and their impact [459]. It typically takes between 30 minutes and 2 hours to complete, depending upon the number of problems reported. The main DAWBA interview is fully structured and has separate sections for all emotional, behavioural and hyperactivity disorders (Table 5.3). The exceptions are that teachers are not asked in detail about emotional disorders and children are not asked about oppositional defiant disorder or hyperactivity. This is because, as discussed in Chapter 2 (p.58), teachers and children do not make good informants for these conditions.

The DAWBA also covers autistic spectrum disorders, eating disorders, tics and psychosis. The sections on these less common disorders were substantially expanded in the B-CAMHS04; in all other respects the DAWBA was identical between the two surveys.

Table 5.3: Emotional, behavioural and hyperactivity disorders covered in the DAWBA

		Parent DAWBA	Teacher DAWBA	Child DAWBA
Emotional	**Separation anxiety**	√		√
	Specific phobia	√		√
	Social phobia	√		√
	Post-traumatic stress disorder [2004 only]	√		√
	Panic	√		√
	Agoraphobia	√		√
	Obsessive compulsive disorder	√		√
	Generalised anxiety disorder	√		√
	Depression	√		√
	'General' section on emotional problems		√	
Behavioural	**Oppositional defiant disorder**	√	√	
	Conduct disorder	√	√	√
Hyperactivity	**Hyperkinetic disorder/Attention Deficit Hyperactivity Disorder**	√	√	

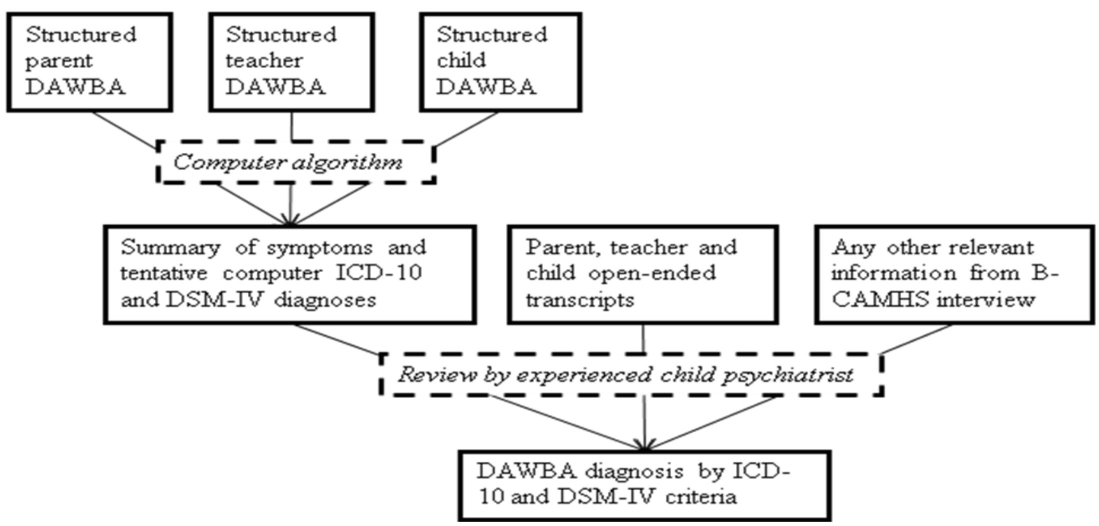
The questions for each disorder closely follow the diagnostic criteria operationalised by DSM-IV and ICD-10, with around 20-25 questions per section [469]. Each section uses skip-rules, however, such that the full set of questions are only administered if children screen positive to initial screening questions (e.g. "Does [Child] ever worry?" for the generalised anxiety disorder section). In addition, children also screened positive for many sections of the DAWBA if they had a high score on the relevant SDQ subscale (e.g. the hyperactivity subscale for the hyperactivity disorder section). This is therefore a conservative skip rule in which the full set of DAWBA questions are administered if either the DAWBA screening questions or the SDQ provide any cause for concern.

Whenever a child did complete a structured section in full, this was followed by open-ended questions. Informants were encouraged to describe the problem in detail and give specific examples, with their answers recorded verbatim by the interviewer. Pre-specified supplementary prompts were used to ensure the respondent's narrative addressed issues of particular clinical concern, such as the frequency, severity and impact of the problem. The interviewers could also make any relevant comments of their own about the respondents or the circumstances of the interview.

Making DAWBA diagnoses

Experienced clinicians then made a diagnosis using a 'case vignette' method [172]. First, as in other fully-structured psychiatric interviews, computer algorithms made tentative DSM-IV and ICD-10 diagnoses based on responses to the fully-structured questions. Trained clinicians then confirmed, overturned or modified the computer-generated diagnosis by reviewing the structured and narrative interview data from across parents, teachers and children (Figure 5.6). This review included careful reading of the transcripts, identifying discrepancies within or between informants, and using the content, length and tone of the transcripts in order to interpret conflicting information [2]. This was done for all children, including those for whom the structured sections provided no evidence of a disorder. These multi-informant clinician-rated diagnoses are henceforth referred to as 'DAWBA diagnoses'.

Figure 5.6: Schematic overview of the process of making DAWBA diagnoses



The B-CAMHS surveys assigned diagnoses using both the ICD-10 and the DSM-IV diagnostic criteria. This provides a choice between the systems when analysing the data, an issue to which I return in Section 6.2.1, Chapter 6. The DAWBA diagnoses were made by a small team of 3-5 child psychiatrists at the Institute of Psychiatry. The lead psychiatrist was always Robert Goodman, who at the time of B-CAMHS99 had 15 years of clinical experience. He trained and supervised the other child psychiatrists, all of whom had completed between three to six years of psychiatric practice after qualifying as doctors. Robert Goodman also discussed any difficult cases with the other psychiatrists, and then independently reviewed all diagnoses or near-diagnoses at the end to ensure that they were rated in a comparable manner. All diagnoses were made blind to the child's ethnicity.

Comparison of the DAWBA with other diagnostic interviews

The DAWBA was designed to bring together the strengths of existing investigator-based and respondent-based interviews. Like fully-structured respondent-based interviews, the DAWBA can be administered by lay interviewers after relatively brief training. Yet like investigator-based interviews, a role is retained for clinical judgement, albeit at the 'overview' rather than the 'interview' stage. This combination of lay interviewers, semi-structured format and clinician-input was novel, and made it economically viable to generate a clinician-rated diagnosis for all 18 415 children in the B-CAMHS surveys. By contrast, previous large epidemiological surveys achieved clinician-rated diagnoses by screening all children with a questionnaire but only administering a diagnostic interview to a subsample [e.g. 433, 470]. In Section 5.2.4 I discussed the considerable advantages of the B-CAMHS one-phase design over this alternative multiphase method.

The main important disadvantage of the DAWBA as compared to traditional investigator-based interviews is the inability of the clinician to seek clarifications or supplementary information. If evidence from the closed questions and transcripts is ambiguous or insufficient, the clinician cannot ask additional questions to get the information they need. On the other hand, assigning diagnoses at the overview stage may make it easier to decide how to interpret disagreement between informants and to decide which children should receive Not Otherwise Specified diagnoses [459]. The requirement for only 3-5 clinicians in B-CAMHS (as compared to 200 lay interviewers) also made it substantially easier to ensure that all diagnoses were made in a comparable manner.

Reliability and validity of the DAWBA diagnoses

The DAWBA has high face validity as a diagnostic tool. As summarised in more detail in Table 14.10, Section 14.2.2 Appendix 2, there is also reasonable evidence for its reliability and validity. Agreement between different psychiatrists rating the same DAWBA cases independently was high in B-CAMHS99 survey for externalising disorders (kappa=0.98) and moderate for internalising disorders (kappa=0.57) [416]. Moreover, these figures may underestimate the consistency and validity of the DAWBA diagnoses in B-CAMHS given that in practice the B-CAMHS clinicians discussed difficult cases with each other and at the end of each survey the lead clinician reviewed all diagnoses.

Unfortunately, little work has been done to validate the DAWBA against other methods of psychiatric assessment. The one small study (N=39) which compared DAWBA diagnoses and case notes diagnoses produced a Kendall *tau* correlation of 0.56 [459]. This is not high, although substantially better than the chance-corrected kappa of 0.21 reported in a meta-analysis of agreement between clinicians and standardised diagnostic interviews [471]. Moreover, it is unclear whether the discrepancies reflect errors from the DAWBA, the case notes, or both. Of course, this problem is to some extent inevitable when no true gold standard exists. The issue is particularly acute in this instance, however, as the authors describe marked inadequacies in many of the case notes.

A paucity of direct comparisons with other diagnostic methods is therefore an important limitation of the existing evidence on the DAWBA. This is, however, partly compensated for by impressive evidence of other forms of construct validity. The DAWBA functions well at differentiating clinic and community samples [459] and produces prevalence estimates [2-3] which are in line with previous literature [4]. Children receiving a DAWBA diagnosis are markedly more likely to present with other indicators of problematic mental health at baseline and show poorer prognosis in terms of both mental health and non-mental health outcomes [459]. The risk factor profiles of children with different types of disorder also differ from the general population and from each other in ways which are in line with the previous literature [3, 417, 459] (see also Section 5.4.3). DAWBA diagnosis in B-CAMHS was also highly predictive of re-diagnosis at three-year follow-up, of mental health service use and of other adverse outcomes such as exclusion from school [48, 445]

Evaluation of the DAWBA outside Britain

The DAWBA has shown good agreement with face-to-face clinician diagnoses in the three settings outside of Britain in which it has been evaluated. In clinic samples in Brazil [469], the Yemen [472] and Bangladesh [466], the DAWBA's sensitivity and specificity relative to face-to-face clinical diagnoses were 89-100% for 'any clinical diagnosis' and 78-100% for particular diagnostic categories. The DAWBA has not been validated in India, but was used with apparent success in a recent survey in Goa [426].

5.3.4 The DAWBA bands

Rationale

The DAWBA provides only a single global assessment of child mental health problems. It also focuses on children at the extreme negative end of the mental health distribution. The parent, teacher and child SDQs complement these limitations by providing informant-specific, dimensional measures. The brevity of the SDQ may, however, render it more prone to bias than the more detailed DAWBA. For example, the SDQ's hyperactivity subscale contains five questions about symptoms while the DAWBA's hyperactivity section contains one screening question, 19 further questions about symptoms, three questions about whether the teacher also reports a problem, one question about age of onset, and six questions about impact. The DAWBA questions are also more specific. For example, while the SDQ asks whether the child is "Restless, overactive, cannot stay still for long", the DAWBA asks "If s/he is rushing about, does s/he find it hard to calm down when someone asks him/her to?" It therefore seems plausible that the DAWBA leaves less scope for cross-cultural differences in norms to create reporting biases. Certainly this seemed to apply to a comparison of British and Norwegian parent and teacher informants which my colleagues and I conducted ([230], Appendix 3).

An informant-specific, ordered categorical variable based upon the DAWBA interview would therefore be a valuable addition to the binary, multi-informant DAWBA and the continuous, informant-specific SDQs. Specifically, it would allow me to investigate whether informant-specific observations in the brief SDQ were also observed in the more detailed DAWBA; to assess exactly which informants contribute to the observed Indian advantage for DAWBA diagnosis; and to increase power when using DAWBA-based outcome measures. Building on work by the original B-CAMHS team, I therefore developed and evaluated a method to create informant-specific, ordered categorical 'DAWBA bands'. Below I outline how I did this: full details are presented in Section 14.2.3, Appendix 2.

Creating the DAWBA bands

Computer-generated predicted probabilities of individual disorders

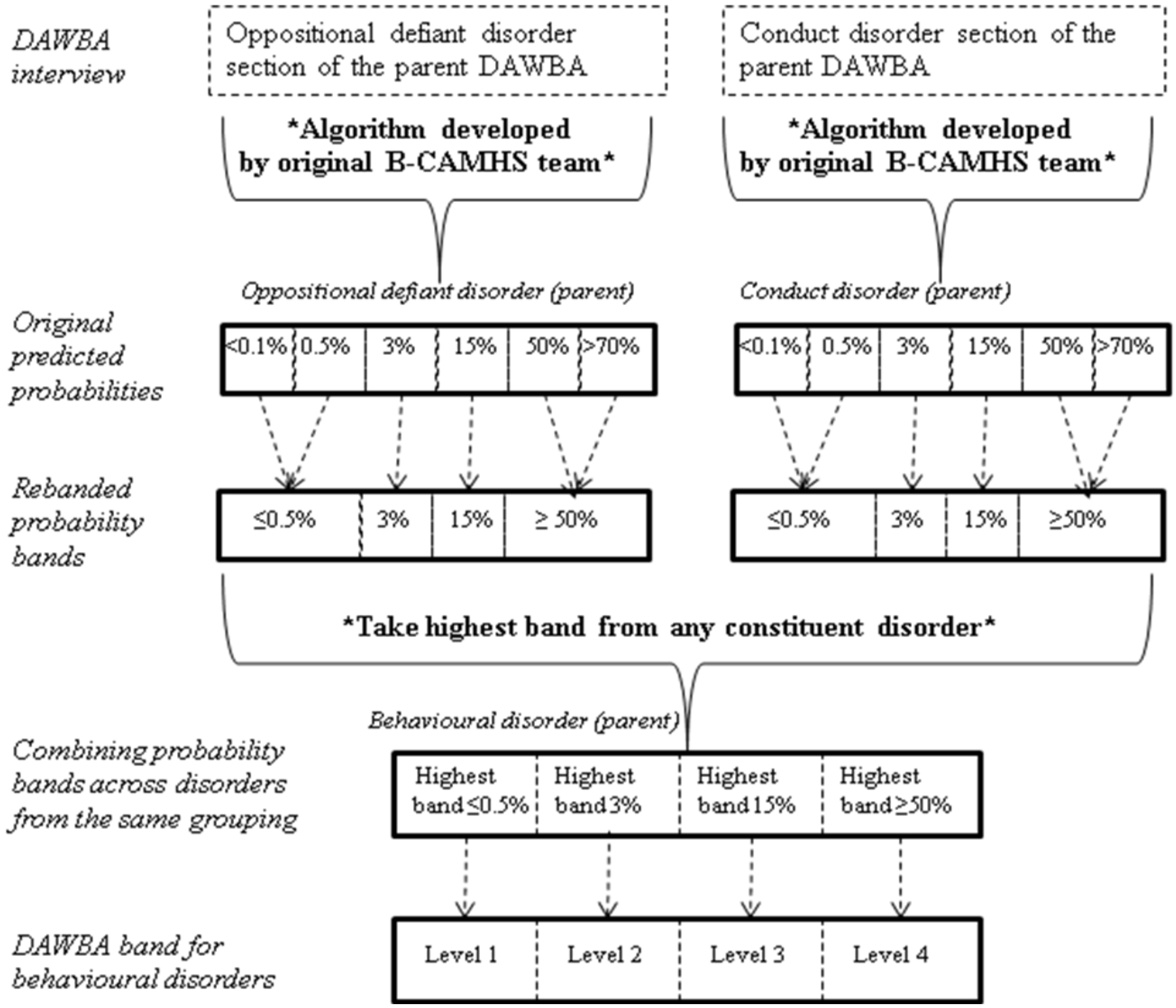
The original B-CAMHS team developed computer algorithms which use the closed sections of the DAWBA to generate informant-specific ordered categorical measures of the probability of receiving a particular disorder. For example, a parent's responses to the structured section on conduct disorder would generate a parent predicted probability score for a diagnosis of conduct disorder. Up to six probability categories were generated for each disorder: <0.1%; 0.5%; 3%; 15%; 50%; >70%. This is illustrated in the top part of Figure 5.7 and described in full in Table 14.11, (Section 14.2.3, Appendix 2).

The algorithms used to create these predicted probabilities were developed empirically using B-CAMHS04 data. This was done through trial-and-error examination of the probability of diagnosis for children with different combinations of symptoms and impact, and also using other criteria such as age of onset where applicable. When applied to B-CAMHS99 the observed probability of disorder generally corresponded closely with the nominal value. For example, among children with a predicted probability of conduct disorder of '50%', about half did in fact receive a DAWBA diagnosis of conduct disorder. The algorithms therefore needed only minor modification before reaching the finalised form used in this PhD (Robert Goodman, personal communication – details not published). Other than this, however, no assessment has been made of the validity of these predicted probabilities.

DAWBA bands for emotional, behavioural and hyperactivity problems

I took these predicted probability scores as my starting point in creating the DAWBA bands. As the top and bottom probability categories usually contained very few children, I first combined the <0.1% and 0.5% categories and combined the 50% and >70% categories. I then created informant-specific variables for 'any emotional disorder', 'any behavioural disorder', 'hyperactivity' and 'any common mental disorder' by taking the highest probability score shown by any of the constituent diagnoses (listed in Table 14.11, Appendix 2). This gave a four point scale of disorder probability from Level 1 at the low end to Level 4 at the high end (Figure 5.7).

Figure 5.7: Schematic illustration of creation of the DAWBA bands



Evaluating the DAWBA bands

I evaluated whether, as intended, the DAWBA bands functioned as ordered-categorical variables of emotional, behavioural and hyperactivity disorders. There was strong evidence that all the DAWBA bands were ordered categorical, with higher DAWBA bands at baseline corresponding to poorer mental health at three-year follow-up (for full results see Table 14.11, Appendix 2). The differences were large at every level, almost always with at least a 2-3 fold increase in prevalence or a mean difference of 2-3 SDQ points. This included mental health outcomes reported by independent informants (e.g. mean teacher SDQ compared to parent DAWBA band).

There was also evidence of good convergent and discriminant validity of the parent, teacher and child DAWBA bands for emotional, behavioural and hyperactivity disorders. In all cases, particular types of DAWBA diagnosis at three-year follow-up were strongly predicted by the corresponding DAWBA band at baseline, and were less strongly or not independently predicted by the other DAWBA bands. For example 'any emotional disorder' at follow-up was strongly predicted by the baseline emotional DAWBA bands (OR>2 per increase in DAWBA band level) and was less strongly or not predicted by the baseline behavioural and hyperactivity DAWBA bands (OR $\leq$1.30)

In combination with their high face validity, I believe this evidence justifies treating the DAWBA bands as informant-specific ordered categorical measures of emotional, behavioural and hyperactivity problems. I therefore use the DAWBA bands in subsequent analyses as a complementary 'intermediate' measures between the multi-informant binary DAWBA diagnosis and the informant-specific dimensional SDQs.

5.3.5 Strengths and limitations of mental health assessment in the B-CAMHS surveys

As argued in Section 2.4, Chapter 2, the most convincing mental health diagnoses combine detailed descriptions of symptoms with an account of how these symptoms are experienced. They also ideally allow some scope for clinical judgement. These principles underpinned the development of the DAWBA for the B-CAMHS surveys. Among the DAWBA's key strengths are its use of explicit, internationally accepted diagnostic criteria (including impact criteria) in generating diagnoses; its use of open-ended transcripts to complement fully structured sections; its use of multiple informants; and its use of clinical judgement in assigning diagnoses. The DAWBA therefore has high face validity as a diagnostic tool, and this is supported by a reasonable amount of evidence as to its reliability and validity in Britain. I therefore consider DAWBA diagnosis to be the single best measure in B-CAMHS of severe and impairing child mental health problems.

The SDQ has several attractive features which complement the DAWBA. The separate SDQ scores from parents, teachers and children allow examination of how far findings such as the Indian mental health advantage are replicated across informants. This provides a basis for identifying potential situation-specific effects and/or cross-cultural biases. The 'external' informant of teachers is particularly useful in this regard. Moreover, the dimensional SDQ offers greater power to detect effects and allows exploration of mental health differences across the full range. This therefore builds upon the concept introduced in Section 2.1.1 Chapter 2 (p.26) of mental health as a continuum. A further useful feature of the SDQ when examining common child mental health problems is that it was designed to contain separate emotional, behavioural and hyperactivity subscales.

Yet as discussed in Section 5.3.2, the SDQ's internal structure has not been appropriately investigated in sufficient detail in Britain. Moreover, neither the DAWBA nor the SDQ have been validated across different ethnic groups in Britain. The studies presented in this Section come from the general UK population, and therefore suggest reliability and validity in the White British majority. It is also reassuring that the DAWBA and SDQ have both shown good psychometric properties in many high-, middle- and low-income settings, including the Indian subcontinent. This does not, however, imply that these measures are reliable or valid in British Indians. I therefore address this issue in Chapter 8, together with an investigation of the more general question of the SDQs internal structure.

Finally, although all previous analyses of B-CAMHS use DAWBA diagnoses and/or SDQ scores as outcomes, I think this does not fully exploit the mental health data collected. Rather I believe these measures could be complemented by an ordered categorical informant-specific measure based upon the structured sections of the DAWBA. For this reason I developed the DAWBA bands and demonstrated that these are ordered categorical mental health outcomes which show good construct validity. In subsequent Chapters I therefore use the DAWBA bands as additional supplementary outcomes.

5.4 Major previous findings from B-CAMHS

Outside of the work reported in this thesis, there have been several important publications resulting from the B-CAMHS studies. Below I describe the major findings to date.

5.4.1 Prevalence of mental disorders

The first aim of the B-CAMHS surveys was to estimate the prevalence of emotional, behavioural and hyperactivity disorders in children in Great Britain. B-CAMHS04 additionally aimed to assess whether these had changed in the previous five years. Table 5.4 shows the prevalence of these disorders using ICD-10 criteria. These figures are essentially unchanged if DSM-IV criteria are used, with the exception that hyperactivity prevalences are around 50% higher.[14]

Table 5.4: Prevalence of common mental disorders by age and sex, in B-CAMHS99 and B-CAMHS04

	Boys (%)		**Girls (%)**		**5-10 year olds (%)**		**11-15/16 year olds (%)**		**All children (%)**	
	1999	**2004**	**1999**	**2004**	**1999**	**2004**	**1999**	**2004**	**1999**	**2004**
Emotional disorders	4.1	3.1	4.5	4.3	3.3	2.4	5.6	5.0	4.3	3.7
Behavioural disorders	7.4	7.5	3.2	3.9	4.6	4.9	6.2	6.6	5.3	5.8
Hyperactivity disorders	2.4	2.6	0.4	0.4	1.5	1.6	1.4	1.4	1.4	1.5
Less common disorders	0.7	1.9	0.4	0.8	0.5	1.3	0.6	1.4	0.5	1.3
Any disorder†	11.4	11.4	7.6	7.8	8.2	7.7	11.2	11.5	9.5	9.6

Sources: Meltzer *et al.* 2000 [2, Table 4.1], Green *et al.* 2005 [3, Table 4.1].
† Sum of columns is greater than the total prevalence because of comorbidity

As Table 5.4 indicates, the overall prevalence of any mental disorder in the B-CAMHS surveys was about 10%. The prevalences in B-CAMHS99 and B-CAMHS04 were very similar for behavioural disorders (around 5.5%), emotional disorders (around 4%) and hyperactivity (around 1.5%). These values are in line with international findings [366] although the proportional morbidity of behavioural disorders is somewhat high. This may reflect the increase in behavioural problems in Britain in recent decades [7]. In the B-CAMHS sample the prevalence of both behavioural and emotional disorders rose with age, while hyperactivity declined somewhat. The prevalence of emotional disorders was slightly higher in girls, while behavioural, hyperactivity and less common disorders were much more prevalent in boys. Again, these findings are in line with other epidemiological surveys [4].

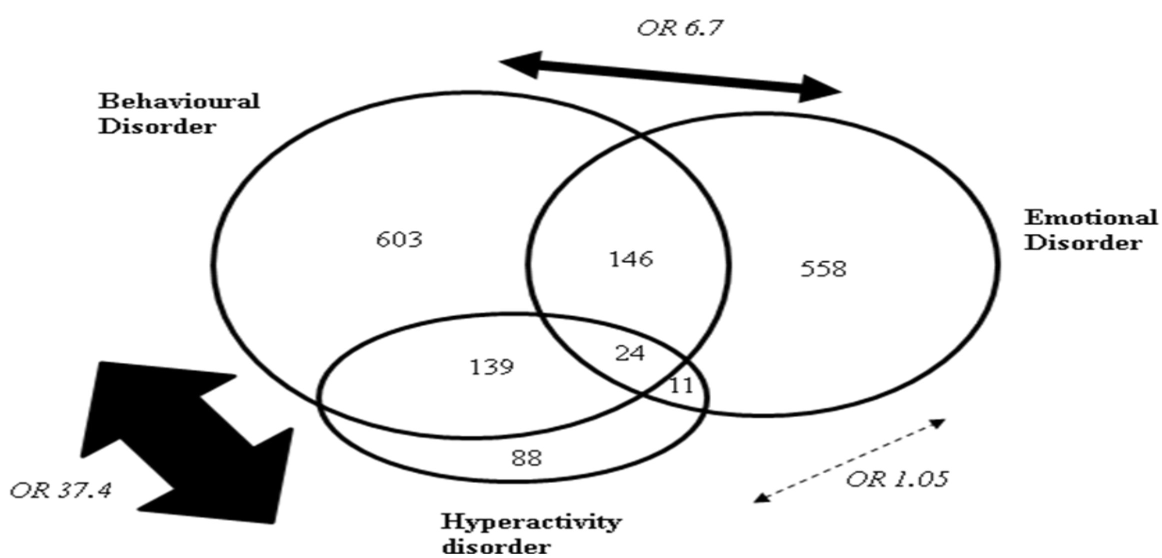
[14] DAWBA diagnoses were assigned in B-CAMHS using both the ICD-10 and the DSM-IV diagnostic criteria. The publications described in this section have been inconsistent in which they use: those produced by ONS use the ICD-10 criteria, while those in scientific journals have mainly used DSM-IV. The relatively close agreement between the two systems (see Box 2.3, Chapter 2, and also Table 6.1, Chapter 6) means this is unlikely to have much effect on the substantive findings.

The only notable change between the two surveys was a marked increase in B-CAMHS04 in the prevalence of less common disorders (0.5% to 1.3%). This was largely driven by an increase autistic spectrum disorders (0.3% to 0.9%). The B-CAMHS research team argues that this is due to the expansion of the 'less common disorder' sections of the DAWBA rather than to any true increase in prevalence [3]. This conclusion is supported by my own comparison of the proportion of children who screened positive on the parent SDQ for autism. This was 0.9% in B-CAMHS99 and 1.1% in B-CAMHS04 (χ^2_1 test p=0.24).

5.4.2 Comorbidity

The B-CAMHS surveys have also confirmed the existence of substantial comorbidity between different disorders using both the DSM-IV [416] and the ICD-10 criteria [2]. Comorbidity was particularly strong between the behavioural and hyperactive disorders, as shown in Figure 5.8.

Figure 5.8: Comorbidity in the common mental disorders in B-CAMHS



Sources: Meltzer *et al.* 2000 [2, Table 4.1], Green *et al.* 2005 [3, Table 4.1]. Numbers give the number of children with each disorder, OR=odds ratios for association between pairs of disorders, adjusting for the third.

5.4.3 Impact and burden of mental disorders

Both B-CAMHS surveys found child mental disorders caused substantial distress and impairment to the children and a substantial burden to their carers. For example, in B-

CAMHS99 children with mental disorders were almost twice as likely to have felt so unhappy or worried that they had asked people for help (41% compared with 23%). Among their parents, 88% reported that their child's problem made them worried, 58% felt it caused them to be depressed and 34% said it made their relationship with their partner more strained [2].

5.4.4 Correlates of child mental disorders

Correlates of 'any mental disorder'

The B-CAMHS publications by ONS present extensive descriptive analyses of the prevalence of mental disorders among children with different characteristics; for example, of different ages, sexes, or from different family types [2-3]. As with overall prevalence rates, the two surveys generally produced very similar findings in terms of both absolute prevalences and relative differences (Table 5.5).

Table 5.5: Family factors associated with a higher prevalence of mental disorders in univariable analyses

Factors associated with a higher prevalence of mental disorders	B-CAMHS99	B-CAMHS04
Lone parent families vs. families with an adult couple	16% vs. 8%	16% vs. 8%
Step families vs. non-step families	15% vs. 9%	14% vs. 9%
Children with four or more resident siblings vs. one resident sibling	18% vs. 8%	11% vs. 9%
Parent had no educational qualifications vs. a degree level qualification	15% vs. 6%	17% vs. 4%
Neither parent working vs. both parents working	20% vs. 8%	20% vs. 8%
Gross weekly household income £0-99 vs. £600 or more	21% vs. 7%	16% vs. 5%
Household reference person's occupation was routine/unskilled (social class V) vs. professional (social class I)	14% vs. 5%	15% vs. 4%
Social sector renting vs. privately renting tenants vs. owner-occupiers	17% vs. 13% vs. 6%	17% vs. 14% vs. 7%

Sources: Meltzer *et al.* 2000 [2, Chapter 4], Green *et al.* 2005 [3, Chapter 4]

The ONS publications also each present one multivariable logistic regression analysis using 'any ICD-10 disorder' as the outcome [2-3]. Additional more extensive multivariable analyses have also been published for the English subsample of B-CAMHS99 [417]. A schematic summary of these findings is presented in Table 5.6. As the many grey-shaded cells indicate, these multivariable analyses have not used all the risk factors available, and this is particularly true for B-CAMHS04. Yet despite this, Table 5.6 demonstrates that child mental disorders are associated with many child, family, area and school factors. The findings are again generally consistent between the two B-CAMHS surveys and also with the previous literature (see Section 2.2.4 Chapter 2). The only two variables which show discrepant findings between the two B-CAMHS surveys are large family size and ethnicity. The latter may reflect the fact that the B-CAMHS99 analyses used only a four-way classification of ethnicity (White, Black, Asian, Other). This may have masked the advantage to Black-African, Indian and Pakistani children seen in B-CAMHS04.

Table 5.6: Factors associated in the published literature with increased risk for any mental disorder in B-CAMHS99 and B-CAMHS04

	Variable	Univariable 1999	Multivariable 1999a	Multivariable 1999b (England)	Univariable 2004	Multivariable 2004
Child	**Older age**	✓	✓	✓	✓	✓
	Male	✓	✓	✓	✓	✓
	Ethnicity (see discussion in text)					✓
	Poor general health	✓		✓	✓	
	Neuro-developmental disorder	✓		✓	✓	
	Low IQ (formal assessment)	✓				
	Poor at reading	✓		✓		
	Severe lack of friends				✓	
	Three or more life events	✓		✓	✓	
	Child frequently punished	✓				
Family	**Lone parent family**	✓	✓	✓	✓	✓
	Step-family	✓		✓	✓	✓
	Parents cohabiting (vs. married)					
	Poor family functioning	✓		✓	✓	
	Poor parent mental health	✓		✓	✓	
	Mother younger when child born	✓			✓	
	Large number of resident siblings	✓	✓			
	Neither parent working	✓	✓		✓	✓
	Rented housing	✓		✓	✓	
	Fewer maternal qualifications	✓			✓	✓
	Low household income	✓			✓	✓
	Parental occupational social class	✓			✓	
Area	**Country of Great Britain**					
	Disadvantaged area (ACORN)	✓	✓		✓	✓
	Deprived area (Carstairs)					
School	**Ford Score for high-risk school**					

Source: Univariable 1999 and multivariable 1999a from Meltzer *et al* 2000 [2]. Multivariable 1999b (England only) from Ford et al. [417]. Univariable and multivariable 2004 from Green *et al.*[3].
✓=variable associated (p<0.05) with a higher rate of 'any disorder'. Variables shaded grey not entered into analyses. B-CAMHS99 classified ethnicity as(White, Black, Asian, Other; B-CAMHS04 classified ethnicity as White, Black Caribbean, Black African, Indian, Pakistani, Bangladeshi, Mixed, Other)

B-CAMHS99 has also been used for more detailed analyses of three risk factors: substance use [473], parenting punishment strategies [474] and accidental injury [475]. These found child mental disorders to be positively associated with substance use, higher parental punishment and accidental injury, but the direction of causality underlying these findings is unclear.

Correlates of emotional, behavioural and hyperactivity disorders

The analysis by Ford *et al.* [417] of the English subset of B-CAMHS99 is the only published analysis of B-CAMHS which models in detail the effects of risk factors on different types of mental disorder. Table 5.7 presents a schematic overview of the associations reported between specific types of disorder and other risk factors. This reveals different patterns of correlates for different disorder categories, including to some extent the two subdomains of emotional disorders (anxiety disorders and depression) and behavioural disorders (oppositional defiant and conduct disorder). Table 5.7 also highlights the fact that different disorders have different *kinds* of profiles. For example, the factors associated with the behavioural disorders are numerous and 'social' while those for hyperactivity are less numerous and mostly biological and cognitive. These findings are in line with the previous literature, as reviewed in Section 2.2.4 Chapter 2. Note, however, that each analysis in Table 5.7 adjusts for comorbidity by entering all other disorder types into the analysis. This may exaggerate the difference in risk factor profiles between disorders because, in the context of high comorbidity, risk factors may only independently predict the disorders with which they are most strongly associated.

Table 5.7: Factors associated with increased risk for subdomains of pathology in the English subset of B-CAMHS99

Risk factor	Anxiety disorder	Depression	ODD	Conduct disorder	ADHD	Any disorder
Older age		✓	✓(protective)	✓		✓
Male			✓	✓	✓	✓
Poor general health	✓	✓				✓
Neuro-developmental disorder	✓		✓		✓	✓
Low IQ			✓	✓		
Poor at reading	✓				✓	✓
Lone parent family	✓					✓
Step-family				✓		✓
Rented housing			✓	✓		✓
Three or more life events	✓	✓				✓
Poor family functioning			✓	✓		✓
Poor parent mental health	✓		✓	✓		✓
Younger mother				✓		
Fewer maternal qualifications	✓			✓		

Source: Ford *et al.* [417].
✓=variable significantly associated ($p<0.05$) with a higher rate of disorder, except older age which was protective for oppositional defiant disorder. For each disorder type, the other types of disorders are included as covariates to adjust for comorbidity.

5.4.5 Persistence, onset and associated risk factors

Finally, three-year follow-up data has been used to calculate rates of persistence and onset for emotional and behavioural disorders. Emotional disorders persisted in 25% of those who had an emotional disorder at baseline in B-CAMHS99 [48] and 30% in B-CAMHS04 [445]. Emotional disorders also developed in 4% of children from B-CAMHS99 without an emotional disorder at baseline and 3% from B-CAMHS04. Persistence and onset for behavioural disorders were 43% and 4% in B-CAMHS99, and 43% and 3% B-CAMHS04.

Multivariable analyses in B-CAMHS99 indicated that the only baseline factor independently predicting persistence of emotional disorder was poor parent mental health, while onset of emotional disorder was predicted by older age, physical illness and stressful life events [48]. Persistence of behavioural disorder was predicted by special educational needs, poor parent mental health, and whether the child was frequently shouted at. Special educational needs and poor parent mental health also independently predicted onset of behavioural disorder, as did male gender and living in a stepfamily. Multivariable analyses were not performed in B-CAMHS04, but the univariable associations with persistence and onset were similar to B-CAMHS99 [445].

Little other work has been done using the B-CAMHS follow-up data. In particular, no use has yet been made of the potential of the longitudinal data to investigate causal directions between child mental health and its correlates. I address this issue in Section 11.1.2 Chapter 11.

5.5 Chapter summary and conclusions

To summarise, the B-CAMHS surveys have many important strengths including the use of a large, representative, population-based samples, and the use of validated mental health measures from multiple informants. The two studies were almost identical in their methodologies, and have produced very similar findings. These findings represent the best available estimates of the prevalence and impact of mental disorders in Britain, and have enhanced and extended what is known about the predictors of child mental disorders. The B-CAMHS surveys have not, however, been used to investigate ethnic differences in any detail, and this includes the apparent Indian mental health advantage. In the next Chapter I describe how I use the B-CAMHS data for this purpose, and begin to present results from my own analyses of the data.

Chapter 6 General methods and preliminary analyses

This section presents a range of important decisions and preliminary analyses regarding the B-CAMHS surveys. First, I describe some general statistical methods which I use throughout this PhD. Note that both here and in subsequent Chapters, Appendix 1 presents a more detailed account of the statistical methods discussed. Second, I outline some key preliminary decisions regarding my choice of mental health outcomes, my approach to combining data between B-CAMHS99 and B-CAMHS04, and my methods for operationalising ethnicity. Finally, I present evidence regarding non-differential participation in B-CAMHS, with particular attention to the possibility of non-differential participation between Indians and Whites.

Together, the decisions and analyses in this Chapter inform and provide a context for the following Chapters in which I address the three central aims of my PhD.

6.1 Data management and statistical analysis

Obtaining data

I obtained the B-CAMHS datasets from colleagues at ONS and the Institute of Psychiatry, London. I did not receive personal identifiers such as the child's name or postcode; instead when using this information for name-matching techniques or assigning area deprivation I visited the ONS offices and performed the matches on site. As described in Chapter 9, I also generated additional variables regarding the schools in B-CAMHS04 in collaboration with the Office for Standards in Education (OFSTED).

Statistical software

I conducted most data analysis using Stata 9.2. The exception was my use of MPlus5 for factor analyses, as indicated in the relevant methods sections.

Adjusting for survey design

As described in Section 5.2.1 Chapter 5, the B-CAMHS surveys sampled children from 901 clusters (postal sectors) in 439 strata. These were sampled without replacement from a total of 8265 postal sectors. The original B-CAMHS team calculated probability weights to account for

1. The oversampling of the Highlands and Islands of Scotland in B-CAMHS99 and the under-sampling of Wales in B-CAMHS04
2. Differential non-response rates by age, sex and region (for details see [2-3])

The original B-CAMHS team additionally calculated three-year follow-up weights, adjusting for the oversampling in B-CAMHS99 of children with disorders at baseline (see Section 5.2.3, Chapter 5).

As discussed more fully in Appendix 1 Section 13.3, adjusting for complex survey design is important when conducting analyses. Failure to weight the data may lead to biased point estimates of means, proportions or effect sizes. It may also bias estimates of variance and standard errors, usually underestimating these in unweighted data. Not adjusting for clustered design is likewise expected to bias (and usually underestimate) estimates of variance. In both cases, this will generate misleadingly narrow confidence intervals, misleadingly large test statistics and misleadingly small p-values. By contrast, failure to adjust for stratification may overestimate the variance, although this effect is often comparatively small.

Both Stata and MPlus have specialised commands for accommodating complex survey design, including stratification, clustering and probability weights. Both estimate parameters using pseudo-maximum likelihood methods and calculating robust standard errors [476-477].

Throughout this PhD, I use these in-built options to adjust for the complex B-CAMHS survey design whenever calculating proportions and means; when fitting regression models (including those using multiple imputation); when conducting exploratory and confirmatory factor analyses; and when calculating Pearson's correlation coefficients. This includes analyses using the three-year follow-up data, which use the follow-up weights. The use of pseudo-maximum likelihood methods means, however, that I cannot adjust for survey design while performing likelihood ratio tests. I therefore instead *compared* models without adjusting for survey design, but then present the better model with adjustment, as follows:

1. Calculate likelihood ratio of nested and general models – *not adjusted for survey design*.
2. Use likelihood ratio to select model – *not adjusted for survey design*.
3. Present results from the chosen model – *adjusted for survey design*.

Furthermore, I do not adjust for survey design when calculating Spearman's coefficients as neither program allows this. In fact the effect of adjusting for survey design was modest in B-CAMHS (see Appendix 1, Table 13.7), meaning that these occasional failures to adjust for survey design are unlikely to affect my substantive findings.

Checking assumptions in regression models

Regression models feature in this and all subsequent data analysis Chapters, with linear and logistic regression being the most common types. Throughout this thesis, I check the assumptions underlying these models as outlined below. Section 13.2 Appendix 1 provides a more detailed discussion of regression techniques in general, and (in Section 13.2.1) of their underlying assumptions in particular.

Linear and logistic models

Assess linearity

- (All models) Plot the outcome (or logit(outcome) for logistic regression) against all continuous or ordered categorical explanatory variables to check for approximate linearity in univariable analysis.
- (All models) Plot the residuals against the expected values to inspect whether these show random scatter around zero.
- (Ordered categorical variables) Likelihood ratio tests to compare linear vs. categorical entry of variables.
- (Continuous variables) Enter quadratic and cubic terms and use the Wald test statistic to determine their significance; or band and enter as categorical.

Normality of the errors:

- (All models) Histograms and normal plots of standardized residuals.

Constant variance of the errors

- (All models) Plot the residuals against the explanatory variable; check that no tendency for the scatter to increase or decrease at higher values.
- Identify influential data points
- (Linear regression models) Sensitivity analyses excluding variables with a Cook's distance of over 4/n.

Dealing with violation of assumptions

Where the relationship between the explanatory and outcome variable was not linear, I entered the variable as an ordered categorical variable or with a quadratic/cubic term.

If the residuals of regression models were skewed rather than normally distributed I repeated the analyses after taking zero-skew logs (see Appendix 1, p.379). I also used these approaches if the variance of the errors was not constant. Both in repeating analyses after taking zero-skew logs and in sensitivity analyses excluding highly influential points, I only report the results of these analyses if there was any substantive difference to the model's findings.

Proportional odds models

Ordered logistic regression requires the proportional odds assumption: that is, that the true population odds ratio for being in category $\geq k$ vs. category $<k$ is the same for all values of k. When using ordered logistic regression, I used likelihood ratio tests to compare the fit of a non-proportional odds model with partial proportional odds model, in which the odds ratios of a given explanatory variable of interest were constrained to be identical. Variables not of substantive interest (e.g. potential confounders such as age or sex) were allowed to have non-proportional odds.

If there was no evidence ($p<0.01$) of a violation of the proportional odds assumption, I selected the partial-proportional odds model; otherwise I selected the fully non-proportional odds model. I used a 1% significance cut-off to reduce spurious findings when fitting multiple models. Having selected the appropriate model, I then reported the results of that model with adjustment for complex survey design.

6.2 Key issues in analysing the B-CAMHS data

6.2.1 Using the DSM-IV diagnostic classification system

DSM-IV and ICD-10 have very similar diagnostic criteria (Box 2.3, Chapter 2). In B-CAMHS, agreement between the two classification systems was perfect or near-perfect for emotional and behavioural disorders (Table 6.1). For hyperactivity disorders the correspondence was poorer, with the narrower ICD-10 definition generating a diagnosis in only 261/408 (63%) of the children who gained a DSM-IV diagnosis.

Table 6.1: Agreement between DSM-IV and ICD-10 diagnoses in B-CAMHS

			DSM-IV	
			No	**Yes**
Emotional disorders	**ICD-10**	**No**	17 406	0
		Yes	0	739
Behavioural disorders	**ICD-10**	**No**	17 483	20
		Yes	0	912
Hyperactivity disorders	**ICD-10**	**No**	18 006	147
		Yes	1	261
Any common disorder	**ICD-10**	**No**	16 755	91
		Yes	0	1569

The overall similarity of DSM-IV and ICD-10 means that the choice of which system to use is unlikely to affect my findings. I have decided to use DSM-IV because the higher prevalence of DSM-IV hyperactivity disorder gives somewhat more power to analyse this rare outcome in Indians. In addition, DSM-IV provides a cleaner division within the behavioural disorders between oppositional defiant disorder and conduct disorder [170] and was used in the most sophisticated existing analysis of B-CAMHS [417].

6.2.2 Pooling data from the two surveys

B-CAMHS99 and B-CAMHS04 were almost identical in their sampling frames, survey design, recruitment methods and mental health assessments (Section 5.2 and 5.3, Chapter 5). The surveys also report very similar disorder prevalences and patterns of correlation with risk factors. I therefore analyse the combined B-CAMHS dataset wherever possible throughout this PhD. This substantially increases power, an important consideration given my focus on a minority group.

Pooling the data does, however, create problems of non-comparability for some explanatory variables. As I discuss in Section 9.2 Chapter 9, some variables were collected at only one time point. In addition, ethnicity data was collected in a slightly different way in the two surveys. I describe how I deal with this in the next section.

6.2.3 Evaluating and extending the B-CAMHS ethnicity data

Combining ethnicity data between B-CAMHS99 and B-CAMHS04

Like many other British surveys (see Section 3.1.2 Chapter 3), B-CAMHS adopted the ethnicity question used in England and Wales in the UK census.[15] This question was revised and extended between the 1991 and 2001 censuses, however, meaning that the two B-CAMHS surveys provided different response options. Table 6.2 presents these response options and my method for combining them. This was the method which maximised the fit between the two censuses in the ONS linked dataset, which links the same 1% of the population across multiple censuses [221, 478]. Of particular interest for this PhD is the consistency in White and Indian ethnicity. For Whites consistency was very high: 99.5% of Whites in 1991 remained in that category in 2001. The corresponding figure for Indians was somewhat lower at 91.0%, indicating a small but non-trivial degree of movement between categories. Among the 'discrepant' Indians, by far the most popular choices of 2001 category was 'Any other Asian background' (49.4%).

[15] The response options are slightly different in the Scottish census question. For example the Scottish census distinguishes 'White Scottish' from 'Other White British' and has 'Black Scottish' as a separate category to 'Black Caribbean' and 'Black African'.

Table 6.2: Eight-category ethnicity classification system

B-CAMHS99 response options (1991 census)	B-CAMHS04 response options (2001 census)	PhD classification†
• White	• White British • Any other White background	• White
• Black – Caribbean	• Black or Black British – Black Caribbean	• Black-Caribbean
• Black – African	• Black or Black British – Black African	• Black African
• Indian	• Asian or Asian British – Indian	• Indian
• Pakistani	• Asian or Asian British – Pakistani	• Pakistani
• Bangladeshi	• Asian or Asian British – Bangladeshi	• Bangladeshi
• Chinese	• Chinese	• Chinese
• Black – Other Black groups • None of these [please describe]	• Black or Black British – Any other Black background • Any other Asian background • Mixed – White and Black Caribbean • Mixed – White and Black African • Mixed – White and Asian • Any other Mixed background • Other ethnic group [please describe]	• Other

† This differs from the method previously used in B-CAMHS04 [3] which placed 'Mixed – White and Black Caribbean' and 'Mixed – White and Black African' in the 'Black' group.

Test-retest reliability in B-CAMHS

Table 6.3 presents the test-retest reliability of the eight-group ethnicity categories between the baseline and follow-up surveys of B-CAMHS99 and B-CAMHS04. Numbers are sometimes too small for meaningful evaluation, but for most groups test-retest reliability appears to be relatively high. This includes Whites and Indians; the proportion of children in the same category at follow-up as at baseline was 99% in Whites (in both surveys) and 90% in Indians (92% in B-CAMHS99, 89% in B-CAMHS04). Among Indians who reported a different ethnicity at follow-up, the most common destination category was 'Other' (8/11), specifically 'Other Asian' (N=3) and 'Mixed White-Asian' (N=5).

These findings are therefore similar to the cross-census comparison in the ONS linked dataset, indicating very high stability in White ethnicity but a small degree of movement to the 'Other' category for Indian children. This movement highlights the fact that ethnic groups are not natural categories which unproblematically accommodate all individuals. Rather, some people may choose different categories in different contexts or according to the options with which they are presented. In the case of Indian ethnicity, however, movement between categories seems to be comparatively rare (≈10%) and is usually to conceptually close categories (e.g. 'Other Asian'). As such, I do not believe this movement represents a fundamental challenge to my use of these ethnicity categories.

Table 6.3: Test-retest reliability of child's ethnicity

		Baseline ethnicity								
		White	**Black Caribbean**	**Black African**	**Indian**	**Pakistani**	**Bangladeshi**	**Chinese**	**Other**	**ALL**
Follow-up ethnicity	**White**	6912 99%	5	3	0	0	0	0	30	6950
	Black Caribbean	7	39 63%	5	0	0	0	0	14	65
	Black African	2	4	41 76%	0	0	0	0	4	51
	Indian	0	0	0	121 90%	8	1	0	10	140
	Pakistani	2	0	0	5	105 90%	0	0	6	118
	Bangladeshi	0	0	0	1	0	32 89%	0	0	33
	Chinese	0	0	0	0	0	0	10 91%	1	11
	Other	51	14	5	8	4	3	1	184 74%	270
	ALL	6974	62	54	135	117	36	11	249	7638

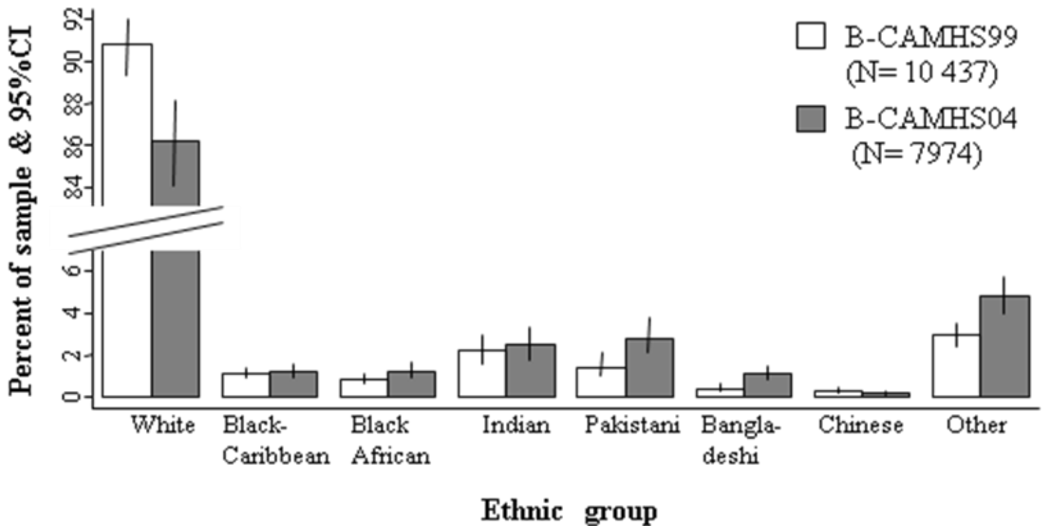
Percentages along the diagonal indicate the proportion of children from each baseline category who were in the same category at follow-up

Distribution of ethnic groups in B-CAMHS99 and B-CAMHS04

The distribution of children across these ethnic groups differed between B-CAMHS99 and B-CAMHS04 (χ^2_1 $p<0.001$). B-CAMHS04 contained more 'Other' children (4.8% vs. 2.9%), and also a somewhat higher proportion of most other minority groups (Figure 6.1). This may partly reflect genuine demographic changes in Britain, given that the proportion of children from most minority ethnic groups is increasing, particularly the 'mixed race' groups [311]. The particularly large increase in the 'Other' group may also reflect its expansion in B-CAMHS04 to contain additional ethnic groups. As for the higher percentage of Pakistani and Bangladeshi participants in B-CAMHS04, I can only speculate that B-CAMHS04 may by chance have included a larger number of postal sectors containing high concentrations of these minority groups.

There are therefore some grounds for concern regarding the comparability of the 'Other' group and the unexplained increase in the proportion of Pakistani and Bangladeshi participants. By contrast, the proportion of children in the Indian group remained fairly constant (2.2% in B-CAMHS99 vs. 2.5% in B-CAMHS04). The change in the White group (90.6% vs. 86.2%) was also not large as a fraction of its total size. This again suggests that for comparisons of these two groups, the change in the classification system did not cause any major disruption.

Figure 6.1: Comparison of ethnic composition of B-CAMHS99 and B-CAMHS04



6.3 Participation in the B-CAMHS surveys

6.3.1 Rationale

This final section of the Chapter assesses the potential for selection bias in B-CAMHS – that is, systematic differences in the characteristics of participants and non-participants. As reported in Section 5.2.1 Chapter 5, the B-CAMHS sampling frame consisted of 26 545 children on the CBC register aged 5-16. 73.2% of these children participated and received a DAWBA diagnosis in B-CAMHS99 and 64.9% in B-CAMHS04. Colleagues at ONS reported that this compared favourably with other national surveys in recent years. Nevertheless these participation rates are not high, and allow considerable scope for selection bias.

When it comes to comparing mental health between groups, differences in the absolute rate of participation (e.g. 70% in one group and 80% in another) will not necessarily be a source of bias if no selection bias exists within each group or if the selection bias is the same in each group. Rather, bias in inter-group comparisons is most likely if the groups differ in the nature or magnitude of selection bias. An example of such a differential selection bias would be if individuals in one group were less likely to take part if they were from deprived areas while a second group showed no such effect. As area deprivation is a risk factor for child mental health problems, this would lead to a misleadingly favourable impression of the mental health of the first group.

For inter-group comparisons the greatest potential for bias therefore results when there is an *interaction* between group membership and some other risk factor with respect to predicting non-participation. Often, however, little is known about the characteristics of non-participants meaning that it is not possible to examine this directly. As such, comparisons of absolute rates may also be informative on the basis that the lower the participation rates the higher the potential for within-group selection bias. Moreover, in the particular case of the B-CAMHS studies, differences between Indians and Whites in absolute rates for teacher and child participation could be an indirect source of *information* bias even if no other selection bias were operating. This is because teacher and child non-participation reduces the amount of information available when making multi-informant DAWBA diagnoses. That this reduces the probability of receiving a diagnosis of mental disorder has previously been demonstrated in B-CAMHS99: children with identical parent-reported mental health were less likely to receive a diagnosis if their teachers did not take part [2, Appendix A].

In this section, I use what limited data is available to me to investigate whether parent, teacher and child participants in B-CAMHS differed systematically from non-participants. I focus upon ethnicity, but also present information on the child's age, sex and area characteristics, and test for interactions between these characteristics and Indian ethnicity. I have previously published a more detailed analysis of non-participation and area deprivation in the English subsample ([479], see Appendix 3).

6.3.2 Methods

<u>Characteristics of participants and non-participants</u>

I defined parent participation as synonymous with overall participation in B-CAMHS. That is, parents were defined as participating if they agreed to take part and if enough information was collected about their child to assign a DAWBA diagnosis. For teachers and children, I defined participation as completing enough of the SDQ to generate a total difficulties score.

I calculated non-participation rates after excluding ineligible children (e.g. children outside the specified age range), in accordance with standard definitions [480]. ONS did not have information on the ethnic group of non-participants, but did receive the child's first and surname for 6115/8142 (75%) of these children. 99% of the remainder were children whose parents opted out in advance to the Child Benefit Centre's (CBC's) first letter. For analyses of parent participation, I therefore imputed ethnicity using the name-matching software Onomap for children whose names were held by ONS. Onomap was developed for use in the UK and has been validated in British adults [481]. As described in Section 14.3 Appendix 2, Onomap shows reasonable predictive properties in B-CAMHS in identifying Indians (sensitivity 71.1%, specificity 99.7%) and Whites (sensitivity 99.1%, specificity 66.6%)..

For analyses of teacher and child non-participation I was able to use the *a priori* more valid method of parent-reported ethnicity. I also compared the results of using parent-reported ethnicity with Onomap ethnicity in teachers and children, in order to assess how far I could have confidence in my analyses of Onomap ethnicity and parent participation.

In addition to the child's name, the CBC also gave ONS the child's date of birth, sex and address. I used this data to investigate selection bias by age, sex, geographical region, metropolitan region and small-area deprivation. I calculated age based upon the child's date of birth and the mid-point of the B-CAMHS fieldwork periods. I used postcodes to assign geographical regions, metropolitan status and Index of Multiple Deprivation (IMD) scores. The IMD is a small-area measure of area deprivation which I describe in detail in Chapter 9 (p.231). The English, Welsh and Scottish IMD differ slightly in their constituent domains and variables, and the raw scores are not directly comparable [482]. I therefore assigned all children an IMD quarter based on their relative position within their country (i.e. England, Scotland or Wales), and then combined quarters across countries. In cases of non-participation due to the family moving without trace, I assigned the IMD score of the child's last known address.

I fitted univariable and multivariable logistic regression analyses with non-participation as the outcome, and tested for interactions between Indian vs. White ethnicity and other predictors of non-response. In doing so I adjusted for the stratification and clustering of the B-CAMHS surveys. I did not use the B-CAMHS weights, however, as these correct for differential response rates by age, sex and geographical region.

Comparison with the general population

It is a limitation of my analyses of parent non-participation that they exclude children whose parents opted out in advance to the CBC and regarding whom ONS received no information. These represent 25% of all non-participants. It is therefore possible that a selection bias might exist which affected only these children. As an indirect assessment of this possibility, I compared the characteristics of the B-CAMHS participants with 1) the age and sex distribution of all children on the CBC register and 2) the age, sex and ethnicity distribution of children in the 2001 UK census.

For ethnicity (but not age and sex), I weighted the B-CAMHS data to adjust for the over and/or under-sampling of Scotland and Wales (see Chapter 5, p.115). This was necessary because the minority ethnic populations of these countries are not representative of Great Britain as a whole. Unfortunately, the only weights provided by ONS adjust simultaneously for non-response by age and sex, and I did not have the necessary data to recalculate weights correcting for regional oversampling alone. I therefore conducted a sensitivity analysis using unweighted figures after restricting the sample to England.

6.3.3 Results

Characteristics of participants and non-participants

Table 6.4 gives the number of participants, non-participants and ineligible informants for parents, teachers and children. Parents and children were ineligible if the child was outside the correct age range and teachers were ineligible if the child was not at school. After excluding those who were ineligible, the non-participation rate was 30.0% (7888/26303) for parents. Among participating parents, non-participation rates were 21.2% for teachers and 10.5% for children.

Table 6.4: Number of participants, non-participants and ineligible individuals among parents, teachers and children

	Parents	Teachers	Children
All Eligible	26303	18263	8577
• Participants	18415	14389	7679
• Non-participants	7888	3874	898
Ineligible	242	152	9838
Overall total	26545	18415	18415

Non-participation by ethnic group

Table 6.5 describes rates of parent, teacher and child non-participation by ethnic group. For parents and teachers, there was strong evidence ($p<0.001$) of ethnic differences in non-participation rates. This was driven by higher rates of non-participation for Black Caribbean, Black African, Pakistani and Other children. By contrast, non-participation was only modestly higher in Indians than in Whites (28.0% vs. 23.5% for parents, $p=0.08$; 25.1% vs. 20.4% in teachers, $p=0.04$). There was no evidence of an overall ethnic difference in child non-participation ($p=0.11$), and rates were very similar between Indians and Whites (11.1% vs. 10.3%, $p=0.70$).

For teachers and children the estimated percentages and substantive conclusions regarding non-participation were usually very similar for parent-reported ethnicity and imputed Onomap ethnicity (Table 6.5). This gives confidence to the results for parent non-participation, for which only Onomap ethnicity is available. The only exception is for Black Caribbeans, a group which Onomap was very poor at identifying (Section 14.3 Appendix 2).

Table 6.5: Parent, teacher and child non-participation by child's ethnicity

	Parents		Teachers				Children			
	Onomap ethnicity		Parent-reported ethnicity		Onomap ethnicity		Parent-reported ethnicity		Onomap ethnicity	
	N	% not participating	N	% not participating	N	% not participating	N	% not participating	N	% not participating
White	22149	23.5%	16315	20.5%	16817	20.6%	7664	10.3%	7895	10.3%
Black Caribbean	16	31.3%	196	28.6%	11	36.4%	91	6.6%	5	0.0%
Black African	192	40.6%	158	35.4%	114	34.2%	71	16.9%	50	10.0%
Indian	500	28.0%	415	25.1%	357	24.7%	207	11.1%	171	14.0%
Pakistani	554	31.2%	372	29.3%	379	30.3%	173	17.9%	172	16.3%
Bangladeshi	118	28.0%	122	27.9%	84	25.0%	56	8.9%	38	7.9%
Chinese	83	33.7%	49	24.5%	55	29.1%	25	8.0%	29	10.3%
Other ethnic group	673	32.8%	632	26.0%	445	28.1%	289	9.7%	217	8.8%
P-value for ethnic differences†		*<0.001*		*<0.001*		*<0.001*		*0.10*		*0.14*
Ethnicity missing	2016	99.95%	4	50.0%	1	100.0%	1	0.0%	0	0.0%
TOTAL††	26303	30.0%	18263	21.2%	18263	21.2%	8577	10.5%	8577	10.5%

†p-value for ethnic difference calculated using univariable logistic regression predicting to non-participation and including all children with non-missing ethnicity data. ††Ineligible individuals are excluded from this table.

Multivariable analyses of non-participation

Table 6.6 presents multivariable logistic regression models predicting to non-participation with ethnicity, age, gender, survey year, geographical region, metropolitan region, and area deprivation as explanatory variables. After adjusting for these additional variables, there was little or no evidence of ethnic variation in non-participation rates in parents (p=0.12), teachers (p=0.06) or children (p=0.36). This included odds of non-participation which were very similar in Indians compared to Whites (OR 1.03 for parents, 1.12 for teachers and 1.02 for children). For both teachers and children, these substantive findings were unchanged when these analyses were repeated using Onomap rather than parent-reported ethnicity. This again adds confidence to the validity of using Onomap for this purpose in the parent analyses.

Regarding the other variables in the model, there was no evidence that age or gender predicted parent non-participation. Teacher non-participation was higher for older children, however, and child non-participation was slightly higher for boys. In all informants, non-participation was higher in B-CAMHS04, in more deprived areas and perhaps in metropolitan areas. The effect of area deprivation was particularly strong, with non-participation increasing across the whole range in an approximately linear fashion (for details, see [479], Appendix 3). There was also evidence of regional variation in parents and children but the nature of this variation was inconsistent across informants. For example, parent non-participation was highest in London and lowest in Wales, while the opposite was true in children. These substantive findings were all identical in univariable analyses examining each of these variables in turn.

Table 6.6: Multivariable model of parent, teacher and child non-participation

		Parent non-participation (OR and 95%CI)	Teacher non-participation (OR and 95%CI)	Child non-participation (OR and 95%CI)
	N	24268	18244	8566
Ethnicity†	**White**	1	1	1
	Black Caribbean	[1.12 (0.37, 3.37)]	1.26 (0.92, 1.72)	0.60 (0.26, 1.35)
	Black African	1.43 (1.01, 2.03)	1.71 (1.22, 2.39)	1.49 (0.74, 2.99)
	Indian	1.03 (0.78, 1.37)	1.12 (0.85, 1.46)	1.02 (0.64, 1.62)
	Pakistani	1.09 (0.88, 1.35)	1.25 (0.94, 1.67)	1.31 (0.81, 2.11)
	Bangladeshi	0.81 (0.50, 1.30)	1.12 (0.71, 1.74)	0.60 (0.21, 1.74)
	Chinese	1.40 (0.91, 2.17)	1.15 (0.58, 2.29)	0.72 (0.14, 3.82)
	Other ethnic group	1.18 (0.99, 1.40)	1.18 (0.95, 1.46)	0.75 (0.51, 1.11)
Gender	**Male**	1	1	1*
	Female	0.98 (0.92, 1.05)	0.95 (0.88, 1.02)	0.84 (0.73, 0.97)
Age	**Change per year**	0.99 (0.99, 1.00)	1.07 (1.06, 1.08)***	1.02 (0.98, 1.07)
Survey year	**1999**	1***	1***	1***
	2004	1.50 (1.40, 1.62)	1.21 (1.09, 1.34)	2.88 (2.40, 3.45)
Geographical region	**South East**	1***	1	1***
	London	1.32 (1.14, 1.52)	1.04 (0.87, 1.25)	0.64 (0.42, 0.99)
	South West	0.81 (0.67, 0.98)	1.00 (0.81, 1.23)	0.84 (0.54, 1.30)
	Eastern	0.89 (0.78, 1.02)	0.99 (0.82, 1.19)	0.57 (0.38, 0.86)
	East Midlands	0.99 (0.82, 1.21)	1.10 (0.90, 1.36)	0.68 (0.45, 1.02)
	West Midlands	0.98 (0.85, 1.15)	1.00 (0.81, 1.24)	0.87 (0.60, 1.25)
	North East	0.89 (0.73, 1.09)	1.02 (0.73, 1.44)	1.60 (1.07, 2.40)
	North West & Merseyside	0.89 (0.77, 1.04)	1.08 (0.89, 1.32)	0.79 (0.56, 1.12)
	Yorkshire & Humberside	0.95 (0.81, 1.10)	1.10 (0.86, 1.40)	0.87 (0.59, 1.28)
	Wales	0.69 (0.56, 0.86)	1.49 (1.11, 1.99)	1.99 (1.41, 2.79)
	Scotland	1.07 (0.90, 1.27)	0.81 (0.64, 1.03)	0.88 (0.60, 1.28)
Metropolitan region	**Non-Metropolitan**	1*	1*	1
	Metropolitan	1.12 (1.01, 1.23)	1.18 (1.03, 1.36)	1.19 (0.95, 1.48)
Area deprivation	**Change per IMD quartile**	1.18 (1.14, 1.22)***	1.15 (1.10, 1.20)***	1.26 (1.17, 1.37)***

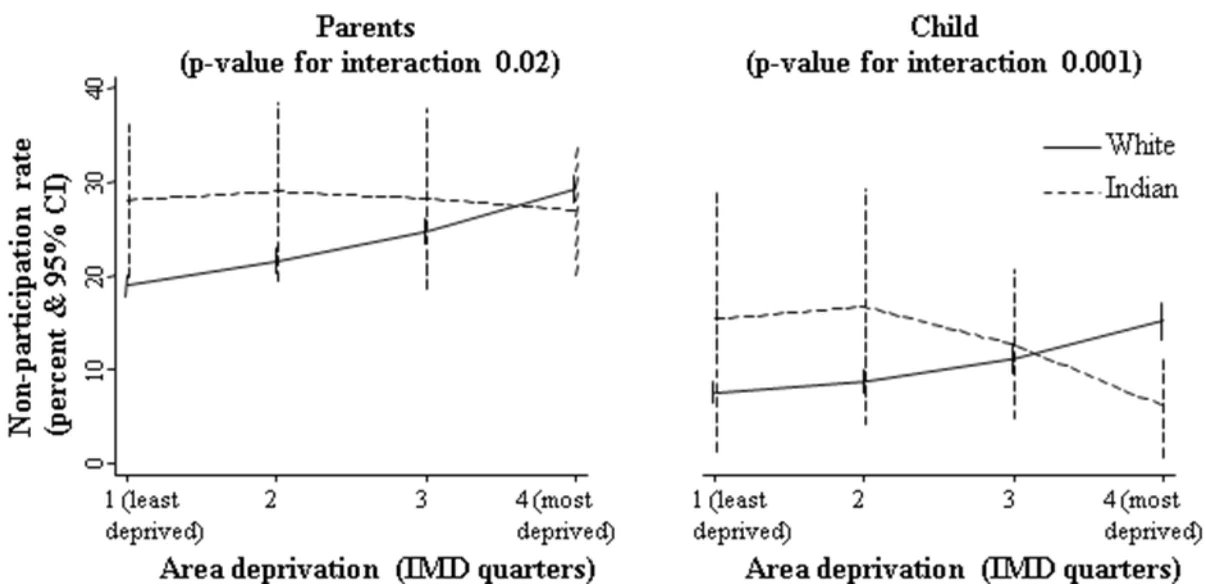
*p<0.05, **p<0.01, ***p<0.001 †Onomap ethnicity for parent non-participation, parent-reported ethnicity for teacher and child non-participation. Note that results are unreliable for Onomap ethnicity for Black Caribbeans.

Interaction between Indian ethnicity and other predictors of non-response

As discussed above, selection biases are particularly likely to jeopardise inter-ethnic comparisons if the nature or magnitude of the biases are different for different ethnic groups. I therefore tested for an interaction between ethnicity and each of the other variables in Table 6.6 after first restricting the sample to Indian and White children.

The only variable for which there was ever any evidence ($p<0.05$) of an interaction was area deprivation. This showed some evidence of an interaction with Indian (vs. White) ethnicity upon parent non-participation ($p=0.02$) and strong evidence for child non-participation ($p=0.001$). In both cases, this interaction reflected the absence of a deprivation gradient for the Indian children, as shown in Figure 6.2. There was no evidence of an interaction between Indian ethnicity and area deprivation upon teacher participation ($p=0.35$).

Figure 6.2: Parent and child non-participation by IMD quartile in Indians and Whites



Comparison with the general population

Table 6.7 presents the age, sex and ethnic composition for 5-15 year olds in B-CAMHS, the CBC register and the 2001 UK census. The distributions are in all cases very similar, including in the proportion of Indians (2.3% in B-CAMHS vs. 2.1% in the census). The results for ethnicity were very similar in a sensitivity analysis restricted to the English subsample of B-CAMHS and not using weights. For example, Indians made up 2.6% of the English B-CAMHS sample and 2.4% of the English census sample.

Table 6.7: Age and sex distributions of children aged 5-15 participating in the B-CAMHS surveys, on the CBC register and in the UK census

		B-CAMHS99 and 04 participants	CBC register, Great Britain, 2000	2001 Census, Great Britain
Age	**5-10**	55.3%	54.7%	54.1%
	11-15	44.7%	45.3%	45.9%
Sex	**Male**	50.7%	51.0%	51.2%
	Female	49.3%	49.0%	48.8%
Parent-reported ethnicity†	**White (White, White Irish and White Other)**	88.9%	[Not available]	88.4%
	Black (Black Caribbean, Black African, Other Black)	2.7%	[Not available]	2.5%
	Indian	2.3%	[Not available]	2.1%
	Pakistani & Bangladeshi	2.7%	[Not available]	3.0%
	Other groups (Chinese, Mixed Race, Other Asian, and Other)	3.4%	[Not available]	4.0%

Source: CBC and Census data from the Office for National statistics [313, 444, 483] and the General Register Office for Scotland [484].
Five-fold ethnicity classification used to achieve comparability between the response options in the English/Welsh and Scottish census questions. Note the B-CAMHS ethnicity figures are weighted for non-response by age, sex and region.

6.3.4 Discussion and conclusions

There was little or no evidence of a difference between Indians and Whites in rates of parent, teacher and child non-participation. The proportion of Indian children was also similar between B-CAMHS and the 2001 UK census. Taken together, these findings provide evidence that participation rates in B-CAMHS were similar for Indians and Whites. This in important in indicating that one of the potential mechanisms for information bias discussed in Section 2.4.2 Chapter 2 does not seem to be an issue in this sample.

By contrast, there was a robust association between greater area deprivation and higher non-participation. This is consistent with many other large, recent surveys in Britain [222, 485-489], although not previously reported for teachers and children in two-phase surveys. Such differential non-participation is important for estimating the prevalence of child mental health problems in the general population, and suggests that B-CAMHS may have underestimated somewhat the prevalence of mental disorders. It also raises the question of whether a differential degree of selection bias with regard to area deprivation could lead to misleading findings when comparing Indians and Whites.

In fact there was some evidence of such a differential selection bias between Indians and Whites. This was, however, in the opposite direction to that necessary to explain the apparent Indian advantage as an artefact. Specifically, the marked gradient in non-participation observed in Whites was absent in Indians. All else being equal, this would be expected to lead to a misleadingly *un*favourable impression of the mental health of Indians relative to Whites.

To summarise, Indian and White parents, teachers and children had similar non-participation rates in B-CAMHS, and the one difference identified between them was an interaction with area deprivation which would be expected to underestimate any Indian advantage. This Chapter therefore provides no evidence that the apparent Indian mental health advantage results from selection bias. This is an important starting point for the next five data analysis Chapters, which address the aims of this thesis to describe and explain the Indian advantage.

Chapter 7 Aim one: Characterising the Indian advantage

7.1.1 Rationale and motivating question

The motivation for this PhD was the estimated prevalence for child mental disorder of 3.4% in Indians as compared with 9.4% in the general population. In this Chapter I address the first aim of my PhD, namely to characterise in detail the nature of this apparent Indian advantage.

I address this aim through simple descriptive analyses comparing Indians and Whites for emotional, behavioural and hyperactivity problems and disorders. I do so using all available mental health measures, namely DAWBA diagnosis; the parent, teacher and child DAWBA bands; and the parent, teacher and child SDQs. Triangulating between these measures is valuable because, as discussed in Section 5.3.5 Chapter 5, each has different strengths and limitations. Specifically, while the multi-informant, clinician-rated DAWBA diagnosis is the best measure of severe mental health problems, it provides only a single global assessment and also focuses only on the negative extreme of the mental health distribution. As an informant-specific and dimensional measure, the SDQ complements these limitations but may be more prone to reporting bias due to its brevity. The DAWBA bands therefore provide an intermediate outcome, being an informant-specific measures based upon the DAWBA. The DAWBA bands also provide a four-point ordinal scale, again being intermediate between the binary DAWBA diagnosis and the continuous SDQ scores.

This Chapter therefore provides an overview of the mental health of Indians and Whites for different types of problem and as reported by different informants. The expectation is that any genuine Indian advantage will be replicated across informants and across outcome measures. Instances where this expectation is met serve to strengthen the evidence for a real Indian advantage, while instances where it is not met highlight important areas for further investigation. These analyses therefore provide the context for the following

Chapter which investigates the validity and comparability of the B-CAMHS mental health measures in Indians and Whites.

7.1.2 Methods

Comparing Indians and Whites

DAWBA diagnoses and DAWBA bands

I first calculated the prevalence DAWBA diagnoses for common child mental disorders in Indians and Whites, and compared the two groups using logistic regression. I then repeated these analyses for the parent, teacher and child DAWBA bands, using ordered logistic regression and assessing the proportional odds assumption as described in Chapter 6, p.150.

SDQ total difficulty scores and subscales

I calculated the mean scores of Indian and White children on the parent, teacher and child SDQ total difficulty score (TDS), and compared the two groups using linear regression. I then repeated these analyses for the separate SDQ subscales. As reviewed in Chapter 5 (p.129), there is some controversy about the correct factor structure of the SDQ with some support for both the hypothesised five-factor solution and an alternative three-factor solution. I evaluate these in Section 8.1 Chapter 8 and therefore present descriptive results for both factor structures in this Chapter.

Age, gender and survey year as a priori confounders

All regression analyses in this Chapter adjust for age, gender and B-CAMHS survey year (1999 vs. 2004) as *a priori* confounders. None of these variables could plausibly generate an Indian advantage in the general UK population, given that Indians and Whites have similar age and sex distributions [313] and there was no large influx of Indian children to Britain between 1999 and 2004. As such, if these variables did explain any of the Indian advantage in B-CAMHS then this would probably reflect chance or selection bias rather than a finding of substantive interest. In fact, Whites and Indians in B-CAMHS were very similar for these variables: the White sample had mean age 10.2 years, was 50.8% male and was 58.0% from B-CAMHS99, while the Indian sample had mean age 10.3 years, was 50.4% male and was 53.0% from B-CAMHS99.

I used likelihood ratio tests to compare models which adjusted for age as a continuous variable and age as a categorical variable (using one-year age bands). These provided no evidence ($p>0.1$) that the categorical model was superior for 'any DAWBA diagnosis' or the parent, teacher or child TDS. I therefore adjust for age as a continuous variable.

In all linear and logistic regression analyses I tested for an interaction between Indian ethnicity and 1) age or 2) gender. Where there was evidence at $p<0.01$ an interaction, I present separate regression models for boys and girls (adjusted for age and survey year) or for 5-8 year olds, 9-12 year olds and 13-16 year olds (adjusted for gender and survey year). I used the more stringent cut-off of $p<0.01$ rather than $p<0.05$ to limit spurious findings from multiple testing.

7.1.3 Results

DAWBA diagnosis

1593/16434 White children and 14/419 Indian children received a diagnosis for any mental disorder. Table 7.1 describes the prevalence of difference types of disorder in both groups, and presents odds ratios for White (vs. Indian ethnicity). These OR are later presented graphically in Figure 7.1 which summarises results for the whole Chapter.

There was strong evidence that Whites were more likely than Indians children to have behavioural disorders (OR 4.39; 95%CI 1.85, 10.38), with similar point estimates for oppositional defiant disorder and conduct disorder (OR 4.49 and 3.70). White children were also more likely to have hyperactivity disorders (OR 9.25; 95%CI 1.29, 66.25), although the rarity of hyperactivity means that the confidence intervals are very wide. By contrast, there was only marginal evidence that White children were more likely to have emotional disorders (OR 1.91; 95%CI 0.93, 3.92), this being driven by some evidence of more anxiety disorders. In no case was there evidence ($p<0.01$) of an interaction between ethnicity and age or gender.

With only 14 Indians with a disorder, any analysis of individual disorders is severely underpowered. Examination of individual diagnoses within disorder subdomains did not, however, provide any indication that particular diagnoses predominated more in Indian than in White children (see Table 14.16, Section 14.4.1, Appendix 2). It was not, for example, the case that all anxiety disorders in Indians were specific phobias.

Table 7.1: Prevalence and odds ratios of disorder types for Indians and Whites

	White (N=16434)		Indian (N=419)		OR & 95%CI for White (vs. Indian) ethnicity
	N	**Prevalence (%) and 95%CI**	**N**	**Prevalence (%) and 95%CI**	
Any disorder	1593	9.63 (9.14, 10.2)	14	3.38 (1.92, 5.87)	3.09 (1.72, 5.52)***
Any common child mental disorder	1516	9.17 (8.69, 9.67)	13	3.14 (1.74, 5.62)	3.14 (1.71, 5.78)***
Any emotional disorder	666	4.02 (3.71, 4.36)	9	2.16 (1.07, 4.32)	1.91 (0.93, 3.92) [p=0.08]
Any anxiety disorder	589	3.55 (3.27, 3.86)	5	1.12 (0.44, 2.86)	3.24 (1.25, 8.43)*
Any depressive disorder	146	0.89 (0.75, 1.06)	6	1.42 (0.62, 3.22)	0.62 (0.26, 1.47)
Any behavioural disorder	858	5.16 (4.80, 5.55)	5	1.24 (0.53, 2.88)	4.39 (1.85, 10.38)**
Oppositional defiant disorder	524	3.13 (2.87, 3.42)	3	0.73 (0.23, 2.23)	4.49 (1.4*3, 14.08)
Conduct disorder	303	1.82 (0.16, 2.06)	2	0.52 (0.13, 1.98)	3.70 (0.93, 14.63) [p=0.06]
Any hyperactivity disorder	386	2.36 (2.13, 2.60)	1	0.26 (0.00, 1.81)	9.25 (1.29, 66.25)*

*p<0.05, **p<0.01, ***p<0.001. OR generated through logistic regression, adjusting for age, gender and survey year.

DAWBA bands

Table 7.2 describes the distribution of the parent, teacher and child DAWBA bands in Indians and Whites. In no case was there evidence at the 1% level that the proportional odds assumption was violated, and Table 7.3 therefore presents the proportional odds for White (vs. Indian) ethnicity upon the DAWBA band. A graphical representation is presented in Figure 7.1, at the end of the results section.

The distribution of all behavioural and hyperactivity DAWBA bands are shifted to the right in Whites compared to Indians, with a smaller proportion of Whites in the lowest risk DAWBA band (Level 1) and a greater proportion in the higher bands. The proportional odds ratios corresponding to these distributions range from 1.54 to 3.18, and were all statistically significant ($p \leq 0.001$ in four out of five cases). By contrast, there was less evidence of an ethnic difference for the emotional DAWBA bands, with only weak evidence ($p=0.03$) that Whites had more problems for the parent and teacher bands and no evidence of a difference on the child band ($p=0.11$).

For none of the DAWBA bands was there evidence ($p<0.01$) of an interaction between ethnicity and either age or gender.

Table 7.2: Distribution of the parent, teacher and child DAWBA bands in Indian and White children

		Parent				Teacher				Child			
		White (N=16334)		Indian (N=368)		White (N=12399)		Indian (N=299)		White (N=6772)		Indian (N=180)	
	DAWBA band	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent	N	Percent
Any common child mental disorder	1 (lowest risk)	4887	30.1%	161	43.5%	7071	57.0%	199	66.8%	2684	39.7%	94	52.5%
	2	8359	51.0%	174	47.2%	3801	30.7%	78	25.9%	2907	42.9%	60	33.2%
	3	1957	12.0%	26	7.3%	865	7.0%	17	5.6%	837	12.3%	18	9.9%
	4 (highest risk)	1131	6.9%	7	2.0%	662	5.3%	5	1.7%	344	5.1%	8	4.4%
Any emotional disorder	1 (lowest risk)	12295	75.3%	295	79.8%	11449	92.3%	286	95.6%	4814	71.2%	140	77.7%
	2	2606	16.0%	51	14.0%	798	6.4%	12	4.0%	1111	16.4%	19	10.7%
	3	954	5.8%	17	4.8%	152	1.25%	1	0.3%	626	9.2%	15	8.3%
	4 (highest risk)	479	2.9%	5	1.4%	–	–	–	–	221	3.2%	6	3.3%
Any behavioural disorder	1 (lowest risk)	6291	38.7%	200	54.5%	7698	62.1%	217	72.9%	3421	50.7%	112	62.7%
	2	8152	49.8%	155	41.9%	3768	30.4%	73	24.1%	2847	41.9%	61	33.4%
	3	1142	7.0%	10	2.8%	337	2.7%	5	1.6%	362	5.3%	4	2.2%
	4 (highest risk)	749	4.6%	3	0.9%	596	4.8%	4	1.4%	142	2.1%	3	1.8%
Any hyperactivity disorder	1 (lowest risk)	14237	87.2%	351	95.4%	9902	79.9%	256	85.7%	–	–	–	–
	2	985	6.0%	12	3.1%	1327	10.7%	25	8.2%	–	–	–	–
	3	804	4.9%	4	1.2%	931	7.5%	16	5.4%	–	–	–	–
	4 (highest risk)	308	1.9%	1	0.3%	239	1.9%	2	0.7%	–	–	–	–

Note that there is no Level 4 teacher DAWBA band for emotional problems and no child DAWBA band for hyperactivity

Table 7.3: Proportional odds ratios for the Indian advantage by DAWBA bands

DAWBA band	Proportional OR for high DAWBA band in White (vs. Indian) children (95%CI)		
	Parent (N=16702)	**Teacher (N=12698)**	**Child (N=6952)**
Any common child mental disorder	1.91 (1.57, 2.32)***	1.55 (1.22, 1.98)***	1.58 (1.19, 2.12)***
Any emotional disorder	1.34 (1.04, 1.73)*	1.84 (1.05, 3.22)*	1.34 (0.94, 1.91)
Any behavioural disorder	2.00 (1.63, 2.45)***	1.69 (1.31, 2.19)***	1.68 (1.24, 2.27)**
Any hyperactivity disorder	3.14 (1.93, 5.13)***	1.53 (1.10, 2.12)*	–

*p<0.05, **p<0.01, ***p<0.001. OR generated through ordered logistic regression, adjusting for age, gender and survey year. Note that there is no child hyperactivity DAWBA band.

SDQ TDS and subscales

Table 7.4 presents the mean TDS and subscale scores for the parent, teacher and child SDQs, and their corresponding regression coefficients – that is, the mean difference in TDS scores between Whites and Indians after adjusting for age, gender and survey year. These are also presented graphically in Figure 7.1.

By parent report, there was strong evidence of an Indian advantage on the behavioural, hyperactivity and prosocial subscales ($p \leq 0.002$). This was partially offset, however, by some evidence of an Indian disadvantage for peer problems ($p=0.02$) resulting in only borderline evidence of an overall Indian advantage for the TDS ($p=0.06$). There was no evidence of any difference between Indians and Whites for the emotional subscale.

The subscales of the teacher and child-report SDQ replicated the results of the parent SDQ in providing strong evidence ($p \leq 0.001$) of substantially fewer problems in Indian children on the behavioural and hyperactivity subscales. Teachers, but not children, also provided evidence of more prosocial behaviour and fewer emotional problems. Neither teacher nor child-report provided any evidence of a difference for peer problems.

All TDS and subscale scores are positively skewed to some extent (see Figure 5.5, Chapter 5). In some analyses, the residuals of the linear regression model were likewise skewed, indicating a violation of one of the assumptions of linear regression. I therefore repeated the analyses using the zero-skew logged TDS scores for linear regression and using ordered logistic regression for the subscales. In all cases the significance levels were similar and the substantive conclusions identical. I therefore present the more readily interpretable linear regression models in Table 7.4.

Interactions with age and gender

Some parent and teacher (but not child) SDQ scales showed evidence of an interaction with age or gender. Specifically, the Indian advantage for parent TDS was observed only in boys (p-value for interaction 0.001). This was driven partly by Indian boys having a greater advantage for the parent hyperactivity scale (p-value for interaction 0.003) and partly by the Indian disadvantage for peer problems being confined to girls (p-value for interaction 0.001). In addition, the teacher-reported Indian advantage for behavioural problems was particularly large in older children (p=0.008 for interaction). Table 14.17 and Table 14.18 (Section 14.4.2, Appendix 2) give details of these interactions in an expanded version of Table 7.4 stratified by age and gender.

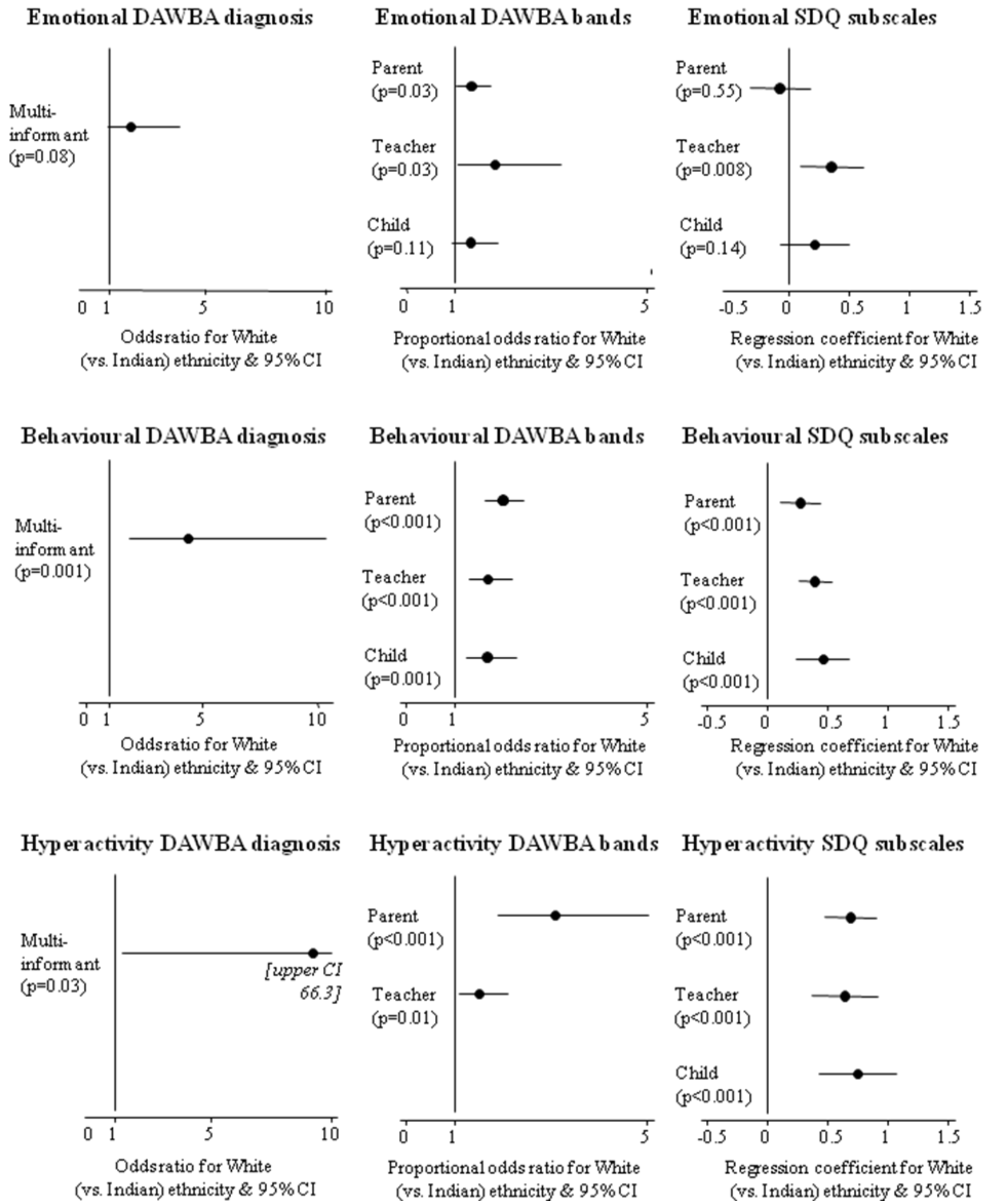
I believe that these apparent interactions should be interpreted with caution. First, they were observed in the context of substantial multiple testing: eight scales/subscales in three informants for interactions with age or gender represents 48 tests in total. Moreover, unlike the other findings in this Chapter, none of these interactions were replicated across informants or on the DAWBA measures. It therefore seems highly plausible that some or all of these interactions represent chance findings due to multiple testing.

Table 7.4: Mean parent, teacher and child SDQ scores for Indians and Whites

	Parent SDQ (N=16386 Whites, 389 Indians)			**Teacher SDQ (N=12796 Whites, 302 Indians)**			**Child SDQ (N=6834 Whites, 183 Indians)**		
	White mean	**Indian mean**	**Regression coefficient & 95%CI for White (vs. Indian) ethnicity**	**White mean**	**Indian mean**	**Regression coefficient & 95%CI for White (vs. Indian) ethnicity**	**White mean**	**Indian mean**	**Regression coefficient & 95%CI for White (vs. Indian) ethnicity**
20-ITEM SCALE									
Total difficulty score (neg)	8.25	7.58	0.63 (-0.03, 1.29)	6.50	5.16	1.34 (0.64, 2.03)***	10.24	8.99	1.26 (0.51, 2.00)**
10-ITEM SCALES									
Internalising subscale (neg)	3.31	3.63	-0.33 (-0.74, 0.09)	2.82	2.52	0.30 (-0.16, 0.76)	4.21	4.15	0.05 (-0.37, 0.47)
Externalising subscale (neg)	4.93	3.95	0.96 (0.64, 1.27)***	3.68	2.64	1.04 (0.69, 1.39)***	6.03	4.85	1.21 (0.73, 1.69)***
5-ITEM SCALES									
Emotional subscale (neg)	1.88	1.96	-0.08 (-0.33, 0.18)	1.49	1.14	0.35 (0.09, 0.62)**	2.73	2.50	0.21 (-0.07, 0.50)
Peer problems subscale (neg)	1.43	1.68	-0.25 (-0.45, -0.04)*	1.34	1.39	-0.06 (-0.30, 0.18)	1.48	1.65	-0.17 (-0.39, 0.05)
Behavioural subscale (neg)	1.56	1.28	0.27 (0.10, 0.44)**	0.89	0.49	0.39 (0.26, 0.53)***	2.19	1.74	0.46 (0.24, 0.68)***
Hyperactivity subscale (neg)	3.38	2.67	0.69 (0.47, 0.90)***	2.79	2.14	0.64 (0.37, 0.92)***	3.84	3.11	0.75 (0.43, 1.07)***
Prosocial subscale (pos)	8.69	8.97	-0.27 (-0.44, -0.10)**	7.41	7.76	-0.34 (-0.55, -0.13)**	7.96	8.08	-0.16 (-0.37, 0.05)

*p<0.05, **p<0.01, ***p<0.001. (pos)=positive scale; higher scores more favourable; (neg)=negative scale; higher scores less favourable. All regression coefficients generated through linear regression, adjusting for age, gender and survey year.

Figure 7.1: Comparison of Whites and Indians for all measures of common mental health problems



7.1.4 Discussion and conclusions

To summarise, the B-CAMHS survey provides strong evidence of a large Indian advantage for behavioural and hyperactivity problems. This is observed with complete consistency across the DAWBA multi-informant clinical diagnoses, the parent, teacher and child DAWBA bands and the parent, teacher and child SDQ subscales. The advantage seems to apply irrespective of age and gender; only for the SDQ was there ever evidence of interactions with ethnicity, and these were not replicated across informants or across measures.

By contrast for emotional problems evidence of a difference between Indians and Whites is weaker and less consistent. There was weak evidence of an Indian advantage for anxiety problems by DAWBA diagnosis and for the parent- and teacher- emotional DAWBA bands (on which, of course, the DAWBA diagnosis depends heavily). For the SDQ, however, the only evidence of an Indian advantage is from teachers, the informant expected to provide the least valid data for this outcome. These results therefore do not provide convincing evidence of an Indian advantage, although certainly they provide no suggestion of a *dis*advantage.

The specificity of the Indian mental health advantage to externalising disorders is consistent with the four non-B-CAMHS studies in my systematic review which disaggregated internalising and externalising problems. In both studies where Indians had an overall advantage this was due to fewer behavioural/hyperactivity problems [377, 391], while in both studies in which Indians had an overall disadvantage this was due to more emotional problems [343, 387]. In B-CAMHS there is no convincing evidence of an Indian disadvantage for internalising problems, the only possible exception being that Indian girls received higher scores on the peer problems subscales of the parent SDQ. This was not replicated in the Health Survey for England 1999, however, the only other large survey in my systematic review to disaggregate the parent SDQ by subscale [343]. The RELACHS study likewise found no difference between Indian and White boys or girls in relationships with friends [419]. In Chapter 8 (Table 8.20) I demonstrate that the anomalous finding in B-CAMHS may reflect a reporting bias among parents who completed the SDQ in translation.

In conclusion, both B-CAMHS and the previous literature provide strong and consistent evidence that the Indian advantage is specific to behavioural and hyperactivity problems. This suggests that comparing Indians and Whites for behavioural and hyperactivity problems is more likely to yield findings of substantive interest than comparisons of emotional problems. It also indicates that it would not be appropriate to use all common mental health problems together as single outcome measure, although it could be appropriate to use a combined 'externalising problems' measure as an outcome.

Yet before accepting the findings of this Chapter, a more rigorous evaluation is needed of the cross-cultural validity of the mental health measures involved. I address this issue in the next Chapter where, in line with my PhD's second aim, I investigate how far the observed differences between Indians and Whites may reflect bias in the measurement of mental health.

Chapter 8 Aim two: Measurement issues in comparing White and Indian children

As reviewed in Section 2.4 Chapter 2, child psychiatry lacks objective and universally applicable methods of classifying and measuring common child mental health problems. Instead it relies heavily upon phenomenological classification and subjective accounts of symptoms and impact. This is an important concern for all psychiatric epidemiology as a potential source of measurement error. The validity of inter-group comparisons can also be undermined through the use of inappropriate mental health constructs in one or more population (a category fallacy) or through measurement biases between populations. Biases may manifest themselves through various potential mechanisms, including systematic differences in thresholds for endorsing closed-question items or in amounts of disclosure to open-ended questions. In turn the source of these systematic differences may be factors such as differences in language of interview or informant type.

When faced with apparent differences between groups, the first question is therefore how far this reflects a real mental health difference and how far it reflects inappropriate or biased measurement. Examining this issue with regard to the apparent Indian advantage is my PhD's second aim, which I investigate by seeking to answer two key questions:

1. Do the common child mental health problems exist in a similar form in Indians and Whites?
2. Is there any evidence of measurement biases which could explain the apparent Indian advantage?

I address these questions in the second and third sections of this Chapter. Before doing so, I provide a further evaluation of the SDQ's internal factor structure in the full B-CAMHS sample. This is necessary because the SDQ subscales are central both for the measurement analyses in this Chapter and for the substantive analyses in subsequent Chapters.

<u>Decision to focus upon both externalising and internalising problems</u>

The evidence presented in Chapter 4 and Chapter 7 suggests that the Indian mental health advantage is specific to behavioural/hyperactive/externalising mental health problems. Nevertheless, throughout this Chapter I evaluate all the SDQ subscales, including the emotional, peer problems and internalising subscales. This is firstly because one of this Chapter's motivating questions is whether the common child mental health problems exist in a similar form in Indians and Whites. One important strategy for assessing this is to examine whether the SDQ items and subscales in Indian children show patterns of correlation which fit with the hypothesised mental health constructs. This requires the inclusion of all SDQ items and scales, not just those for externalising problems. Secondly, a similar prevalence of emotional/internalising problems in Indians and Whites might itself be a finding of substantive interest – but only if it were shown to be a real similarity based on measures with comparable psychometric properties in the two groups. Finally, even if externalising mental health problems formed the focus of subsequent substantive analyses, I would still wish to include internalising problems as one potential explanatory variable. This again requires me to establish the validity of these measures.

8.1 Evaluation of the SDQ's factor structure

8.1.1 Rationale and motivating questions

As discussed in Chapter 5 (p.129), the SDQ's internal factor structure has not been investigated as thoroughly as would be desirable. In particular, the international literature provides some suggestions that the proposed distinction between emotional, peer problems, behavioural and hyperactive problems may not be justified, and that instead these should be combined into an internalising (emotional plus peer problems) and an externalising (behavioural plus hyperactivity) subscale. Table 14.7 (Section 14.2 Appendix 2) summarises the constituent items of these different subscales.

In this section I evaluate the SDQ's factor structure, focusing on two main questions:

- **Internal factor structure of the SDQ**: Which of the proposed SDQ factor structures shows the best fit to the B-CAMHS data?

- **Construct validity of SDQ factors**: Which of the proposed SDQ subscales show convergent and discriminant validity? I address this in two ways, looking both at construct validity across informants and as compared to DAWBA diagnosis.

8.1.2 Methods

This section introduces several statistical techniques which I use at various points in this and subsequent Chapters. Appendix 1 provides a fuller description of their underlying methods and assumptions, and a fuller justification my choice of these techniques over alternatives.

Internal factor structure of the SDQ

When a hypothesised factor structure exists for a set of items, a model-based framework such as confirmatory factor analysis (CFA) provides a means of testing of whether a proposed factor structure fits the data. It also allows formal comparisons of the relative fit of competing factor structures [490].

Model specification

Box 8.1 summarises the models which I assessed for the parent, teacher and child SDQs in the full B-CAMHS dataset (i.e. analysing all ethnic groups together). I first assessed a five-factor second order model in which the emotional, peer, behavioural and hyperactivity subscales formed part of the total difficulty score (TDS) and the prosocial score was separate (Figure 8.1). This is what Mellor and Stokes [464] argue corresponds to the hypothesised SDQ factor structure in their thorough assessment of the SDQ in an Australian sample. If this model did not show adequate fit to the B-CAMHS data, I evaluated the additional three- and five- factor models described in Box 8.1 and illustrated in Figure 8.2 and Figure 8.3.

One rationale for these analyses is that the best-fitting model in the full B-CAMHS sample can be a starting point for multi-group analyses comparing Indians and Whites. A further important rationale is to investigate whether the emotional, peer, behavioural and hyperactivity scales can be treated as separate constructs, or whether I should combine them into internalising and externalising problems. In some ways, the inclusion of the items on the prosocial scale complicates this second issue. I therefore conducted a further set of CFA analyses which excluded the five prosocial items (Box 8.1; Figure 8.4 and Figure 8.5).

Box 8.1: CFA Models fitted for parent, teacher and child SDQs

Hypothesised SDQ structure (Figure 8.1).
Factors
- TDS/emotional/peer/behavioural/hyperactivity/prosocial

Model structure:
- A five-factor second order model with the emotional, peer, behavioural and hyperactivity subscales as part of the TDS and the prosocial score separate

Additional three-/five-factor models (Figure 8.2 and Figure 8.3).
Factors:
- Internalising/externalising/prosocial
- Emotional/peer/behavioural/hyperactivity/prosocial

Model structures
- A first order model
- A general-specific model in which the 'general' factor included all 25 SDQ items.[16]
- A second order model in which the second order factor included all first order factors.

Two-/four-factor models (Figure 8.4 and Figure 8.5).
Factors:
- Internalising/externalising
- Emotional/peer/behavioural/hyperactivity

Model structures:
- A first order model
- A general-specific model in which the 'general' factor included all 20 items
- A second order model in which the second order factor included all first order factors; not fitted for two-factor model because it had too few known parameters to be freely estimated.

Model estimation

I performed the CFA in MPlus5, using a multivariate probit analysis [491] with the extension for ordinal data [492] and estimating model fit using the Weighted Least Squares, mean and variance adjusted (WLSMV) estimator. I included in my analyses all parent, teacher and child SDQs with scores for all SDQ subscales. A small number of individual items were missing for these individuals (<0.7% for all informants in both Indians and Whites).[17] For these I used MPlus's default 'pairwise present' estimation option, which estimates parameters based on pairs of items in turn using all individuals with observations

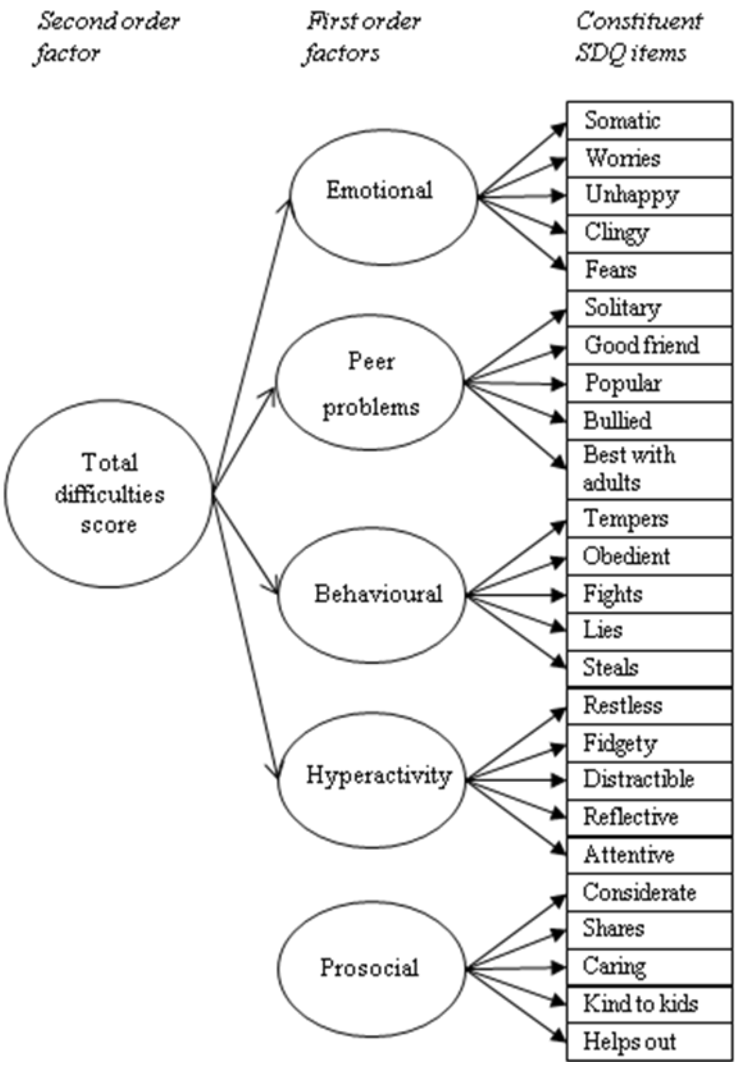
[16] I also fitted general-specific models in which the 'general' factor included only the 20 items of the TDS, but these invariably showed substantially worse fit than models with the 25-item general factor.
[17] The original B-CAMHS team calculated subscale scores where one or two of the five items were missing by scaling up and then rounding the score of the three or four observed items. For example, a score of 7 from four observed items would be scaled up to (7/4)*5=8.75, then rounded to 9.

for any given pair. See Section 14.5.1 Appendix 2 for an example of the MPlus syntax I used, corresponding to the hypothesised SDQ factor structure for the parent SDQ.

Assessing goodness of fit

I follow common practice in reporting multiple indices of fit, namely the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI) and the Root Mean Square Error of Approximation (RMSEA) [490, 493]. To consider a model as showing acceptable fit, I required a CFI>0.90 and ideally >0.95; TLI>0.90 and ideally >0.95; and RMSEA<0.08 and ideally <0.05 [490]. I also checked that the standardised loadings of each variable onto its factor(s) were of reasonable magnitude (>0.4).

Figure 8.1: Five-factor second-order model corresponding to the hypothesised SDQ structure[18]



[18] See Section 14.5.1 Appendix 2 for corresponding MPlus syntax

Figure 8.2: Additional five-factor models

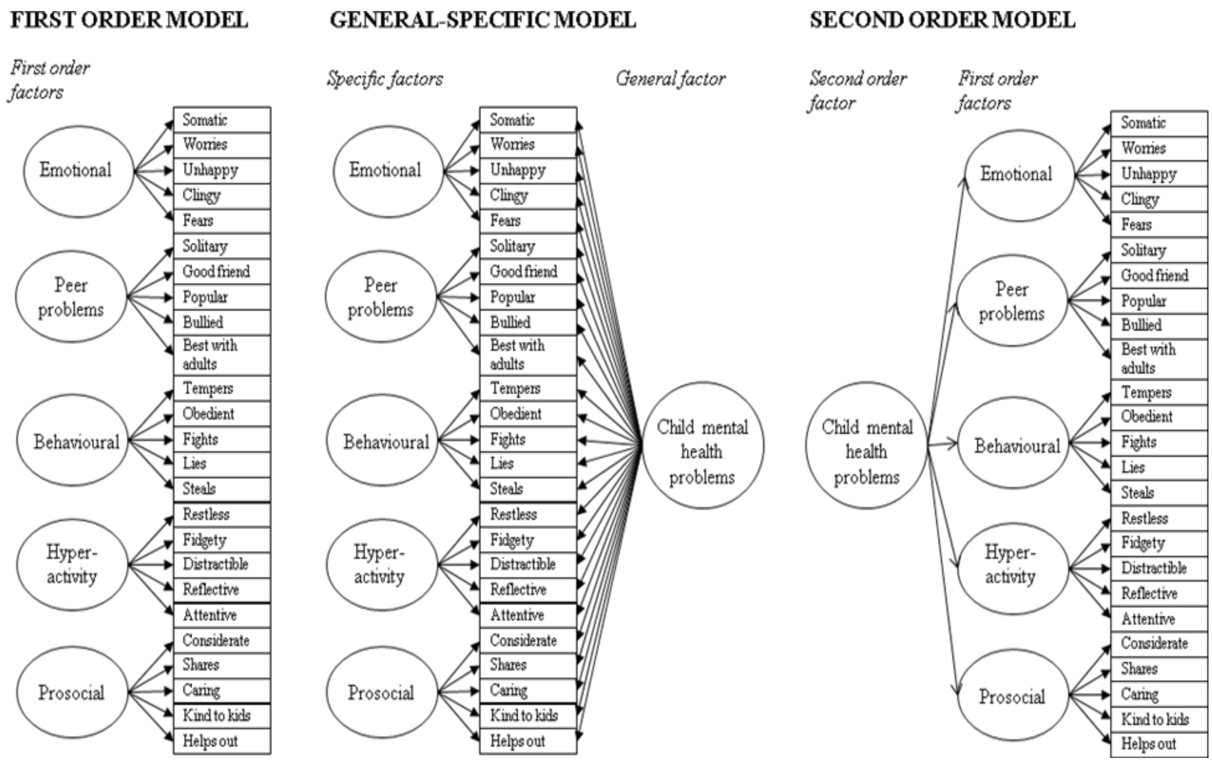


Figure 8.3: Three factor models

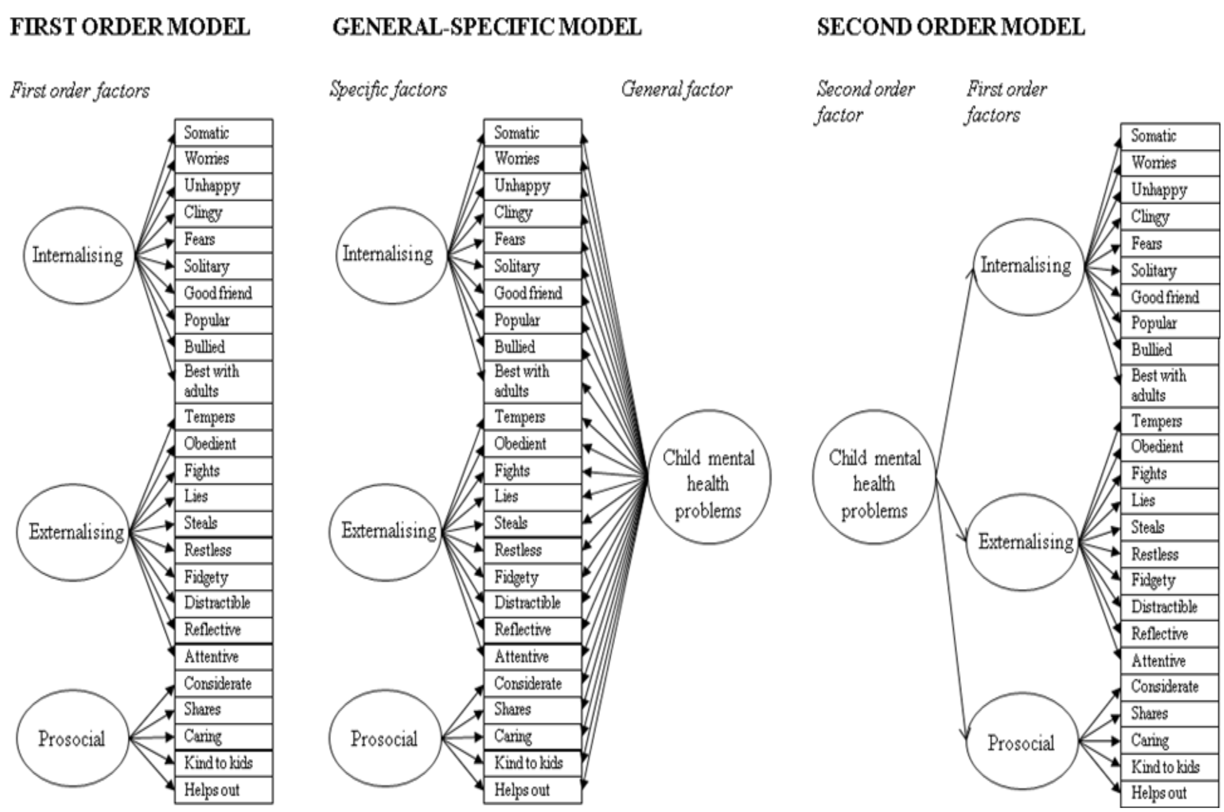


Figure 8.4: Four factor models

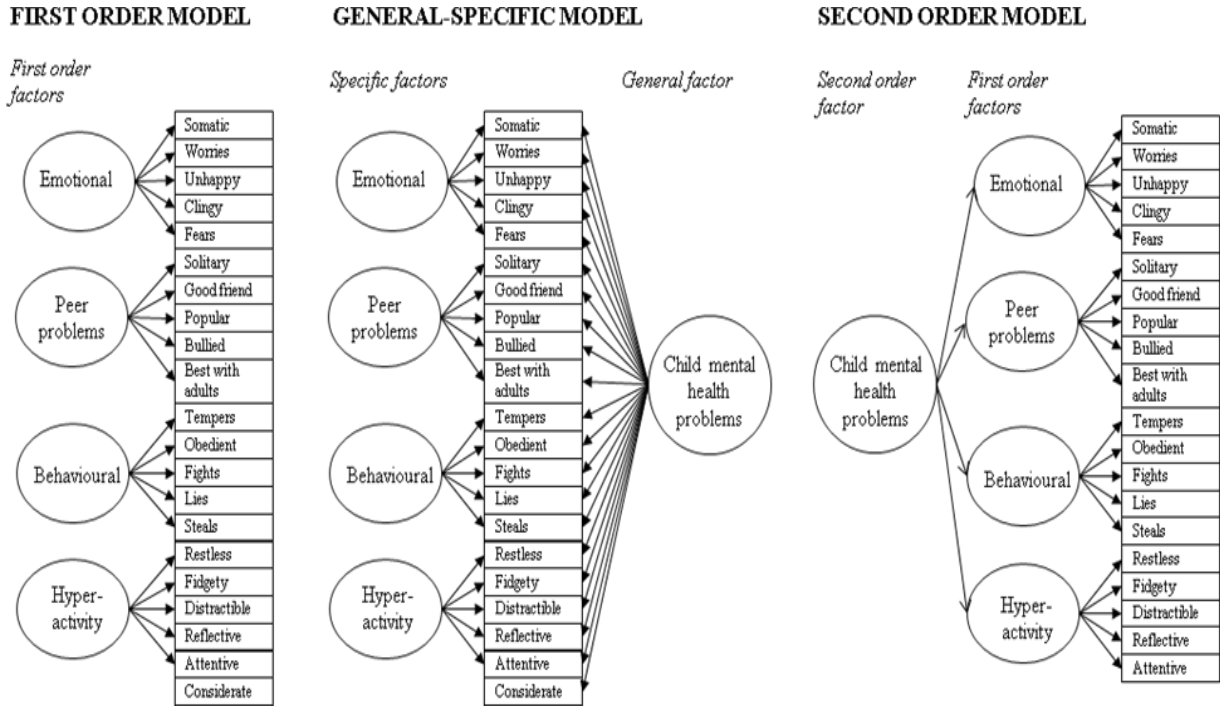
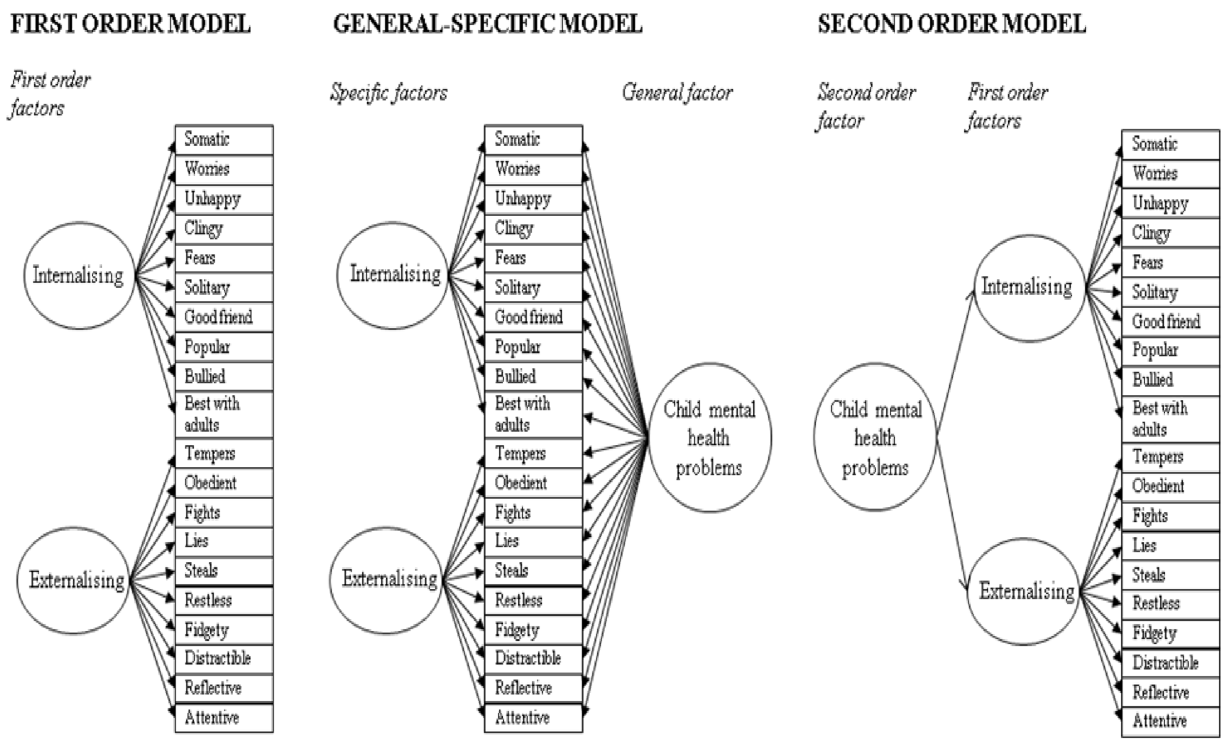


Figure 8.5: Two factor models



Construct validity of the SDQ

Construct validity of the SDQ subscales across informants

Multitrait-multimethod (MTMM) analyses are a method for assessing the construct validity of a set of measures [494-495]. As illustrated in Table 8.1, MTMM are based on a correlation matrix of multiple traits (e.g. the proposed SDQ subscales) as measured by multiple methods (e.g. by parent and teacher). Together, this provides evidence on several aspects of convergent and discriminant validity:[19]

1. **Test-retest reliability for each trait** ([a] cells in Table 8.1). This could not be assessed in B-CAMHS because the necessary 'retest' was not carried out.
2. **Within-method correlation of different traits ([b] cells).** Good construct validity requires within-method correlations between traits which are consistent with *a priori* hypotheses. For example, if prosocial behaviour is hypothesised to be more closely related to externalising than internalising problems, then the prosocial-externalising coefficients should be larger than the prosocial-internalising coefficients within all informants.
3. **Between-method correlation of the same trait ([c] cells or validity coefficients, shaded grey).** Good convergent validity requires high correlation between measures of the same trait assessed by means of different methods (e.g. parent externalising score and teacher externalising score). In addition, the magnitude of the [c]-cells should ideally be larger than the [b] cells. If not then this indicates that the 'method factor' (i.e. the informant) is a more powerful determinant of a child's score than the 'trait factor' (i.e. the child's mental health).
4. **Between-method correlation of different traits ([d] cells).** Good discriminant validity requires that the correlation between different traits be substantially lower than the correlation between the same traits. So, for example, the parent externalising score should be substantially less highly correlated with the teacher internalising and prosocial scores than with the teacher externalising score.

[19] As described in Section 13.1 Appendix 1, convergent validity requires highly correlation with other measures of the same or similar traits, while discriminant validity requires little or no correlation with measures of different traits.

Table 8.1: Schematic representation of an MTMM analysis for the parent and teacher SDQ

		Parent			Teacher		
		Internalising	**Externalising**	**Prosocial**	**Internalising**	**Externalising**	**Prosocial**
Parent	**Internalising**	*[a]*					
	Externalising	*[b]*	*[a]*				
	Prosocial	*[b]*	*[b]*	*[a]*			
Teacher	**Internalising**	*[c]*	*[d]*	*[d]*	*[a]*		
	Externalising	*[d]*	*[c]*	*[d]*	*[b]*	*[a]*	
	Prosocial	*[d]*	*[d]*	*[c]*	*[b]*	*[b]*	*[a]*

Cells labelled [a] show the test retest reliability coefficients for each subscale for each informant. Cells labelled [b] show agreement between different subscales reported by the same informant. Cells labelled [c] are the validity coefficients, and show correlations between the same subscales reported by different informants. Cells labelled [d] show agreement between different subscales reported by different informants. Together the [c] and [d] cells form a heterotrait block.

I used MTMM analyses to evaluate the construct validity of the five- and three-factor SDQ factor structures, calculating the MTMM coefficients using Spearman's correlations. I used Spearman's correlations rather than intra-class correlations (ICCs) because the ICC's appropriate for analysis of the B-CAMHS data [ICC (1,1)[20]] are influenced both by the *consistency* between measures (e.g. whether children with high parent-reported scores also receive high teacher-reported scores) and by differences in *absolute* means (e.g. if parent scores are on average higher than teacher scores) [496]. This property of ICCs is often useful, but in this instance it would not be of interest if mean parent SDQ scores were higher or lower than mean teacher scores.

I calculated MTMM coefficients using all individuals with data for the SDQ(s) in question. So, for example, I calculated the within-parent [c] coefficients using all children who had parent SDQs and calculated the parent-teacher [a] and [b] coefficients using all children with parent and teacher SDQs. The MTMM analyses are therefore based shifting subsets of children. This is clearly not ideal but I felt it was preferable to restricting my analyses to the 5684 children with SDQs from all three informants. These 5684 children represent only 32% of all children with parent SDQs, 41% of those with teacher SDQs and 76% of those

[20] This is the ICC in which different children are rated by different individuals, and in which the analysis is of individual scores not mean ratings; See Appendix 1, p.359.

with child SDQs. They are also a non-random subset, most notably for only including children aged 11 or over, but also because of the other participation biases described in Section 6.3 Chapter 6.

Construct validity of the SDQ subscales relative to the DAWBA

MTMM analyses assess the SDQs construct validity by making comparisons across informants. Comparisons between the SDQ and the DAWBA provide a further opportunity to assess construct validity. The *a priori* prediction is for DAWBA diagnoses of emotional disorders to correlate most closely with the emotional SDQ subscale of the parent, teacher and child SDQs; behavioural disorders with the behavioural subscale; hyperactivity with the hyperactivity subscale; and autistic spectrum disorders with the peer problems and prosocial subscales.

To test whether this was the case, I performed a series of logistic regression analyses, with DAWBA diagnoses as outcomes and the SDQ subscales for a particular informant as explanatory variables.[21] The four outcomes I used were DAWBA diagnosis for any emotional disorder, any behavioural disorder, any hyperactivity disorder, or autistic spectrum disorder. For the explanatory variables, I used both the five-factor and the three-factor SDQ subscales.

Predicting baseline DAWBA diagnoses using baseline SDQ subscale scores would be somewhat circular because, as described in Chapter 5 (p.124), the SDQ subscales form part of the skip rules for some DAWBA sections. I therefore capitalised upon the inclusion in B-CAMHS of a three-year follow-up and used DAWBA diagnoses at follow-up as outcomes. I decided *a priori* not to use child-reported SDQs to predict autistic spectrum disorder because only 10/71 children with a follow-up diagnosis of autism completed the child interview, and these children may lack insight as informants.

[21] This therefore adapts the approach used in Chapter 5 (p.138) to assess the construct validity of the DAWBA bands.

8.1.3 Results

Internal factor structure of the SDQ

Table 8.2 presents the CFAs fitting the hypothesised second order five-factor model to the parent, teacher and child SDQs. Defining acceptable fit as CFI>0.9 (ideally >0.95), TLI>0.9 (ideally >0.95) and RMSEA<0.08 (ideally <0.05), it can be seen that for no informant did this model provide acceptable fit by all three indices. Moreover, at least one index of fit was always substantially below the acceptability cut-off (CFI in parents and children, RMSEA in teachers). Minor model modifications such as allowing correlation between the residual variances of some items did not achieve acceptable fit in any informant.

Table 8.2: Summary of CFAs for the hypothesised SDQ factor structure

Informant	Factors	Model structure	CFI	TLI	RMSEA
Parent (N=18222)	Emo/peer/behav/hyp/pro	2nd order	0.850	0.915	0.065
Teacher (N=14263)	Emo/peer/behav/hyp/pro	2nd order	0.901	0.952	0.097
Child (N=7678)	Emo/peer/behav/hyp/pro	2nd order	0. 801	0.846	0.072

Emo=emotional, peer=peer problems, behav=behavioural, hyp=hyperactivity, pro=prosocial.

I therefore proceeded to fit additional first order, general-specific and second order models with both three and five factors. Of the 18 models (three model structures times two factor structures times three informants), only the three-factor general-specific model for the teacher SDQ showed acceptable or near-acceptable fit by all three indices (Table 14.19, Section 14.5.2 Appendix 2). After restricting the analyses to the 20 TDS items, however, acceptable fit was shown for 8/9 indices (three informants times three indices) for both a two-factor general-specific model and a four-factor first order model (Table 8.3; full results in Table 14.20, Section 14.5.2, Appendix 2). This therefore provides some support both for treating the emotional, peer, behavioural and hyperactivity scales as separate constructs or alternatively for analysing them simply as internalising and externalising scales. It further indicates that the prosocial subscale may be problematic in its relation to the other variables.

Table 8.3: Selected models from additional CFAs on the 20 TDS items

Informant	Factors	Model	CFI	TLI	RMSEA	Acceptable fit by all indices
Parent (N=18222)	Int/ext	General-specific	0.897	0.937	0.062	(√)
	Emo/peer/behav/hyp	First order	0.907	0.945	0.058	√
Teacher (N=14263)	Int/ext	General-specific	0.950	0.974	0.072	√
	Emo/peer/behav/hyp	First order	0.919	0.962	0.086	(√)
Child (N=7678)	Int/ext	General-specific	0.909	0.928	0.053	√
	Emo/peer/behav/hyp	First order	0.900	0.924	0.054	√

Int=internalising, ext=externalising, emo=emotional, peer=peer problems, behav=behavioural, hyp=hyperactivity. Acceptable fit defined as CFI and TLI>0.9, RMSEA<0.08.

For all the above CFAs, I checked that individual items had acceptable loadings (>0.4) onto their hypothesised factors. The majority of items showed acceptable loadings in all models, this being particularly true of the externalising items. For example, in the two-factor internalising/externalising general-specific model, all ten externalising items in all three informants had loadings of >0.4 on the general and/or their specific factor. By contrast, a few internalising items were borderline with ‘somatic’ and ‘best with adults’ showing maximum loadings of 0.3-0.4 in parents and ‘good friend’, ‘popular’ and ‘best with adults’ showing maximum loadings of 0.3-0.4 in children.

Construct validity of the SDQ subscales

Construct validity of the SDQ subscales across informants

Table 8.4 presents an MTMM analysis of the five SDQ subscales. All informants showed a similar pattern of within-method correlation of the five subscales, and one which was consistent with *a priori* expectations. Thus the two ‘externalising’ behavioural and hyperactivity subscales were always most closely associated with each other, and the resulting correlation was the largest observed in the triangle. The two ‘internalising’ emotional and peer subscales were likewise usually most closely associated with each other, although the teacher peer problems and child emotional subscales were also highly correlated with other subscales. The prosocial subscale was always much more highly correlated with the behaviour and hyperactivity subscales than with the emotional or peer problems subscales.

Turning to the heterotrait blocks, the validity coefficients (shaded in grey) were all significantly different from zero. The absolute values were low, however, ranging from

0.29-0.47 for the behavioural and hyperactivity subscales and from 0.20-0.36 for the emotional and peer subscales. This is similar to the magnitude of the within-informant correlation of different subscales, indicating that the informant effect on SDQ scores is at least as strong as the trait effect.

Low inter-informant correlations and high informant effects are not desirable measurement properties, but are unfortunately typical for child mental health questionnaires (see Chapter 2, p.60). More worrying is the failure of some subscales to show cross-method discriminant validity. In particular:

- In all three informant pairs, behavioural disorders did not show good discriminant validity relative to hyperactivity problems (relevant cells circled with solid line). For example, in the parent-teacher block the parent behavioural-teacher behavioural correlation is 0.31 which is no higher than the parent behavioural-teacher hyperactivity correlation (0.31) and slightly lower than the parent hyperactivity-teacher behavioural correlation (0.33).
- The teacher prosocial subscale did not show discriminant validity relative to the behavioural and hyperactivity subscales reported by either parents or children (relevant cells circled with dashed line). For example, the teacher prosocial subscale correlated 0.25 with the parent prosocial scale, -0.25 with the parent behavioural scale and -0.28 with the parent hyperactivity subscale.

These results therefore indicate that although the five SDQ subscales show a pattern of within-informant inter-correlation in line with *a priori* expectations, cross-informant comparisons do not show good discriminant validity for the behavioural, hyperactivity and prosocial subscales. These analyses therefore do not support the claim that these subscales all tap into *distinct* aspects of child mental health problems in this population. When the MTMM was repeated for the internalising, externalising and prosocial subscales, convergent and divergent validity was more satisfactory for the internalising-externalising contrast (see Table 8.5). The prosocial scale still showed poor discriminant validity relative to the externalising scale, however, particularly in the case of the teacher prosocial score.

Table 8.4: MTMM analyses for the five SDQ subscales

		Parent					Teacher					Child				
		Emo	**Peer**	**Behav**	**Hyp**	**Pro**	**Emo**	**Peer**	**Behav**	**Hyp**	**Pro**	**Emo**	**Peer**	**Behav**	**Hyp**	**Pro**
Parent	**Emo**															
	Peer	0.37														
	Behav	0.29	0.28													
	Hyp	0.26	0.26	0.49												
	Pro	-0.12	-0.17	-0.40	-0.32											
Teacher	**Emo**	0.24	0.20	0.12	0.14	-0.05										
	Peer	0.14	0.28	0.17	0.19	-0.13	0.41									
	Behav	0.03	0.15	0.31	0.33	-0.18	0.18	0.36								
	Hyp	0.07	0.17	0.31	0.47	-0.19	0.25	0.33	0.60							
	Pro	-0.05	-0.15	-0.25	-0.28	0.25	-0.16	-0.40	-0.56	-0.57						
Child	**Emo**	0.36	0.20	0.14	0.13	-0.02	0.20	0.14	0.03	0.04	0.00					
	Peer	0.19	0.34	0.12	0.13	-0.05	0.16	0.22	0.08	0.08	-0.08	0.32				
	Behav	0.19	0.15	0.42	0.37	-0.25	0.11	0.14	0.29	0.30	-0.24	0.33	0.21			
	Hyp	0.15	0.09	0.27	0.40	-0.17	0.13	0.09	0.25	0.33	-0.22	0.32	0.17	0.52		
	Pro	-0.03	-0.08	-0.16	-0.17	0.30	-0.01	-0.09	-0.14	-0.16	0.24	-0.03	-0.15	-0.32	-0.30	

Emo=emotional SDQ subscale, peer=peer problems, behav=behavioural, hyp=hyperactivity, pro=prosocial. N=18222 parents; N=14263 teachers and N=7678 children. N=14139 for the parent-teacher comparison, N=7561 for the parent-child comparison and N=5755 for the teacher-child comparison. Values in cells are Spearman's correlation coefficients. Heterotrait blocks are outlined in bold, validity coefficients are shaded grey. Cells circled with solid lines indicate problematic discriminant validity for the behavioural subscale relative to the hyperactivity subscale. Cells circled with dashed lines indicate problematic discriminant validity for the prosocial subscale relative to the behavioural and hyperactivity subscales

Table 8.5: MTMM analyses for the internalising and externalising SDQ subscales

		Parent			Teacher			Child		
		Int	**Ext**	**Pro**	**Int**	**Ext**	**Pro**	**Int**	**Ext**	**Pro**
Parent	**Int**									
	Ext	0.37								
	Pro	-0.18	-0.40							
Teacher	**Int**	0.30	0.22	-0.11						
	Ext	0.14	0.48	-0.21	0.36					
	Pro	-0.11	-0.31	0.25	-0.32	-0.62				
Child	**Int**	0.40	0.18	-0.04	0.25	0.08	-0.04			
	Ext	0.20	0.48	-0.23	0.15	0.37	-0.26	0.37		
	Pro	-0.07	-0.19	0.30	-0.06	-0.16	0.24	-0.09	-0.35	

Int=internalising, ext=externalising, pro=prosocial SDQ subscales. Number of participants as in Table 8.4.

Construct validity of the SDQ subscales relative to the DAWBA

Table 8.6 shows which of the five SDQ subscales at baseline independently predicted DAWBA diagnoses at three-year follow-up. For the parent and teacher SDQ, the expected subscale(s) were the strongest predictor for each type of DAWBA diagnosis. This included autistic spectrum disorder, which was most strongly predicted by the peer problems and prosocial subscales. For the child SDQ the evidence of discriminant validity was less convincing: the emotional subscale was no more strongly associated with emotional disorder than the peer problems subscale, and the hyperactivity subscale was less strongly associated with hyperactivity disorder than the behavioural subscale.

The five-factor structure therefore showed convergent and discriminant analyses for parent and teacher SDQ but not for the child SDQ. By contrast, the three-factor structure showed convergent and discriminant analyses for all three informants (Table 8.7).

Table 8.6: Independent association of the five SDQ subscales at baseline with DAWBA diagnosis at follow-up (OR and 95%CI)

		Emotional DAWBA diagnosis	Behavioural DAWBA diagnosis	Hyperactivity DAWBA diagnosis	Autism DAWBA diagnosis
Parents (N=7856)	**Emotion**	1.32 (1.25, 1.39)***	1.00 (0.94, 1.06)	0.95 (0.86, 1.04)	1.25 (1.08, 1.44)**
	Peer problems	1.14 (1.06, 1.22)***	1.08 (1.01, 1.16)*	1.28 (1.15, 1.43)***	1.59 (1.37, 1.84)***
	Behavioural	1.16 (1.08, 1.24)***	1.64 (1.54, 1.76)***	1.34 (1.22, 1.48)***	0.65 (0.55, 0.76)***
	Hyperactivity	1.00 (0.95, 1.05)	1.22 (1.16, 1.29)***	1.76 (1.58, 1.96)***	1.43 (1.24, 1.64)***
	Prosocial	1.07 (1.00, 1.16)	0.97 (0.91, 1.03)	1.15 (1.03, 1.27)*	0.55 (0.48, 0.62)***
Teachers (N=6173)	**Emotion**	1.16 (1.09, 1.23)***	0.98 (0.93, 1.04)	0.92 (0.83, 1.02)	1.16 (1.02, 1.30)*
	Peer problems	1.10 (1.02, 1.19)**	1.10 (1.03, 1.18)**	1.25 (1.12, 1.38)***	1.38 (1.20, 1.58)***
	Behavioural	1.11 (1.02, 1.22)*	1.31 (1.22, 1.40)***	1.11 (1.01, 1.23)*	0.81 (0.66, 0.99)*
	Hyperactivity	1.01 (0.95, 1.07)	1.19 (1.13, 1.25)***	1.51 (1.38, 1.66)***	1.22 (1.07, 1.40)**
	Prosocial	0.99 (0.93, 1.06)	0.97 (0.90, 1.04)	1.01 (0.91, 1.12)	0.70 (0.59, 0.83)***
Child (N=3283)	**Emotion**	1.24 (1.14, 1.35)***	0.98 (0.89, 1.07)	0.87 (0.71, 1.07)	–
	Peer problems	1.27 (1.15, 1.40)***	1.04 (0.93, 1.16)	1.27 (0.96, 1.68)	–
	Behavioural	1.04 (0.92, 1.16)	1.64 (1.48, 1.81)***	1.49 (1.16, 1.90)**	–
	Hyperactivity	1.06 (0.98, 1.15)	1.13 (1.04, 1.23)**	1.29 (1.06, 1.58)*	–
	Prosocial	1.11 (1.00, 1.22)*	1.03 (0.93, 1.14)	1.01 (0.83, 1.22)	–

*p<0.05, **p<0.01, ***p<0.001. Subscales expected *a priori* to be the strongest predictors are shaded grey. Autistic spectrum disorder was not used as an outcome for the child SDQ.

Table 8.7: Independent association of the three SDQ subscales at baseline with DAWBA diagnosis at follow-up (OR and 95%CI)

		Emotional DAWBA diagnosis	Behavioural DAWBA diagnosis	Hyperactivity DAWBA diagnosis	Autism DAWBA diagnosis
Parents (N=7856)	**Internalising**	1.24 (1.20, 1.27)***	1.04 (1.00, 1.07)	1.08 (1.02, 1.15)**	1.42 (1.34, 1.50)***
	Externalising	1.06 (1.02, 1.09)**	1.38 (1.33, 1.42)	1.54 (1.45, 1.64)***	1.00 (0.94, 1.06)
	Prosocial	1.07 (0.99, 1.14)	0.93 (0.88, 0.99)	1.15 (1.04, 1.27)***	0.58 (0.51, 0.65)***
Teachers (N=6173)	**Internalising**	1.14 (1.10, 1.18)***	1.04 (1.00, 1.07)*	1.05 (0.99, 1.10)	1.25 (1.17, 1.33)***
	Externalising	1.05 (1.01, 1.08)*	1.24 (1.20, 1.28)***	1.33 (1.26, 1.40)***	1.02 (0.95, 1.09)
	Prosocial	0.99 (0.94, 1.06)	0.95 (0.89, 1.02)	1.00 (0.90, 1.11)	0.69 (0.58, 0.83)***
Child (N=3283)	**Internalising**	1.25 (1.19, 1.31)***	1.02 (0.96, 1.08)	1.03 (0.89, 1.19)	–
	Externalising	1.05 (1.00, 1.10)*	1.32 (1.25, 1.40)***	1.36 (1.20, 1.54)***	–
	Prosocial	1.11 (1.00, 1.22)*	1.00 (0.90, 1.11)	0.95 (0.78, 1.16)	–

*p<0.05, **p<0.01, ***p<0.001. See notes to Table 8.6.

8.1.4 Discussion and conclusion: choosing the internalising and externalising SDQ subscales as primary outcomes

The motivating question for this section is whether the emotional, behavioural and hyperactivity subscales can be analysed separately, or whether they should be combined into the broader internalising and externalising subscales. These analyses provide some support for both positions. Confirmatory factor analyses demonstrated acceptable fit in all informants from models with both two- and four-factor structures once the prosocial items were excluded. The internalising and externalising subscales showed a cleaner pattern of convergent and discriminant validity across informants in the MTMM analyses, and there were several instances where cross-informant discriminant validity between the behavioural and hyperactivity subscales was not achieved. Yet distinguishing between the emotional, peer, behavioural and hyperactivity subscales of the parent and teacher SDQs did seem to add value when predicting to DAWBA diagnosis. Likewise, the prosocial scale showed poor discriminant validity in MTMM analyses relative to the externalising subscales, but better construct validity in predicting autistic disorders.

There is thus a discrepancy between the MTMM analyses and the analyses using the SDQ to predict DAWBA diagnosis. One possible explanation is that the MTMM analyses reflect patterns of subscale association in the *full* B-CAMHS sample, which is mostly comprised of children without mental health problems. In this low-risk, general population sample there may not always be a clear-cut distinction between behavioural and hyperactivity symptoms or between externalising symptoms and prosocial behaviour. By contrast, discriminating symptom clusters may be easier when predicting to DAWBA diagnoses and so focusing on high-risk children. An analogy from clinical practice would be the greater ease of distinguishing depressive and anxiety disorders in mental health specialist clinics than in the general population [173].

I therefore conclude that there may be no single best SDQ factor solution, but rather that this may depend in part upon one's study population and study aims. For this PhD, I believe that the internalising and externalising subscales are the most appropriate candidate outcomes. This decision is motivated by three considerations. First, this thesis focuses upon a group with an apparent mental health advantage. This group is therefore expected to show a low-risk rather than a high-risk pattern of association. Second, as described in Chapter 7, the Indian advantage is consistently observed for both the behavioural and hyperactivity subscales but not for the emotional and peer subscales. I believe this justifies combining these four subscales into the externalising and internalising subscales. Finally, a smaller number of scales with more items each increases power for psychometric and substantive analyses – an important consideration given the comparatively small Indian sample in B-CAMHS.

Throughout the remainder of this PhD, I shall therefore compare Indians and Whites in terms of the internalising and externalising subscales. This decision provides the starting point for next section, in which I evaluate the psychometric properties of these subscales in Indians and Whites.

8.2 Cross-cultural comparability of mental health problems in Indians and Whites

8.2.1 Rationale and motivating questions

In this section I examine whether the common child mental health problems exist in a similar form in Indians and Whites. Specifically, and as justified below, I address the following questions:

1. Do individual symptoms of common child mental health symptoms show a similar pattern of interrelation in Indians and Whites?
2. Is the construct validity of the proposed mental health constructs comparable between Indians and Whites?
3. Do mental health symptoms have the same implications for impact and burden in Indians as in Whites?

Symptom interrelation and construct validity

An etic/universalist approach assumes that the same mental health problems and disorders exist in all populations (although prevalence may vary). As argued in Section 2.3.3 Chapter 2, the literature suggests that the common child mental health problems can indeed often be observed cross-culturally. I know of no work which investigates this issue in British Indian children, however, or indeed in Indian children living in India or other parts of the world. As such, a central issue in this PhD is whether the common child mental health problems are sufficiently similar in Indians and Whites to allow meaningful comparisons using the DAWBA and SDQ.

In this section I address this by investigating whether the SDQ internal factor structure is similar in Indians and Whites; that is, whether the internalising and externalising symptoms ‘hang together’ as expected in both groups. I then evaluate whether the SDQ subscales show a similar degree of construct validity across informants. This analysis is particularly informative because it includes teachers, a culturally ‘external’ informant. Finally, I compare the SDQ’s construct validity relative to the DAWBA in Indians and Whites.

Of course, if the SDQ does *not* show a similar internal factor structure or comparable construct validity, then it will not be easy to distinguish between a problem with the underlying constructs (e.g. 'externalising problems' is not a meaningful construct in Indians) and a problem with the measures (e.g. 'externalising problems' is a meaningful construct, but the SDQ is not a valid measure of them in Indians). For the purpose of informing subsequent sections of this PhD, however, this arguably does not matter – if the SDQ were invalid in Indians for whatever reason, then it would not be appropriate to use it as a mental health outcome in substantive analyses.

Association between symptoms and impact/burden

In addition to investigating the factor structure and construct validity of the SDQ symptoms, I also examine their implications for burden and impact. Ethnic differences in this respect could reflect either methodological issues (e.g. differences in the validity of the measures) or substantive differences (e.g. ethnic differences in how far problems are exacerbated or accommodated). Nevertheless, the centrality of impact to the definition of mental health problems means that at a minimum higher symptom counts should be associated with greater impact and burden. If this were not observed then it would undermine the assertion that the mental health constructs measured by the SDQ truly are *problems* for the children involved.

This therefore represents one form of hypothesis testing analysis. One could also test other hypotheses by, for example, comparing risk factor associations in Indians and Whites. I decided not to pursue this route, however, because the disorder-specific associations which *a priori* seem most likely to be cross-culturally invariant involve rare risk factors – for example, hyperactivity and neuro-developmental disorder, behavioural problems and school expulsion, emotional problems and deliberate self-harm. By contrast, for many common risk factors it would not be straightforward to interpret an ethnic difference in the association between the risk factor and mental health. For example, if socio-economic position were not associated with externalising problems in Indian children then this would not necessarily imply that externalising problems were not a meaningful construct; it might instead reflect ethnic differences in the meaning of socio-economic indices such as 'education' or an interaction such that the negative effects of socio-economic disadvantage were weaker in Indians.

A focus upon the SDQ

The analyses in this section focus upon the SDQ for three reasons. First, all SDQ items were asked of all informants. This makes it straightforward to analyse the internal structure of symptoms on the SDQ and to compare symptom and impact/burden scores. By contrast, most DAWBA items were only asked of the minority of children who screened positive for a particular section. Secondly, the SDQ is a dimensional measure which has been validated separately in parents, teachers and children. This makes it straightforward to examine the construct validity of the SDQ across different informants. By contrast, the only informant-specific DAWBA measures are the DAWBA bands, which have only a four-point range and which have not been extensively validated. Finally, the dimensional nature of the SDQ also means that it offers greater power for comparing Indians and Whites. The SDQ therefore represents a particularly attractive potential outcome for my subsequent substantive analyses, and this makes it all the more important to provide a thorough investigation of its cross-cultural validity in this Chapter.

8.2.2 Methods

<u>Internal factor structure of the SDQ in Indians and Whites</u>

Exploratory factor analyses

Confirmatory factor analyses (CFAs) are more appropriate than exploratory factor analyses (EFAs) if one has a prior hypothesis regarding factor structure. I felt, however, that the transparency of the empirically-derived EFA factors would make them a useful descriptive starting point when comparing the SDQ factor structure between Indians and Whites. I therefore conducted EFAs on the 20 TDS items, specifying two factors and running the analyses separately in Indians and Whites. I fitted the EFAs in MPlus5 using the extension for ordinal data. I rotated the factor loadings using an oblique geomin rotation and report all item loadings >0.3.

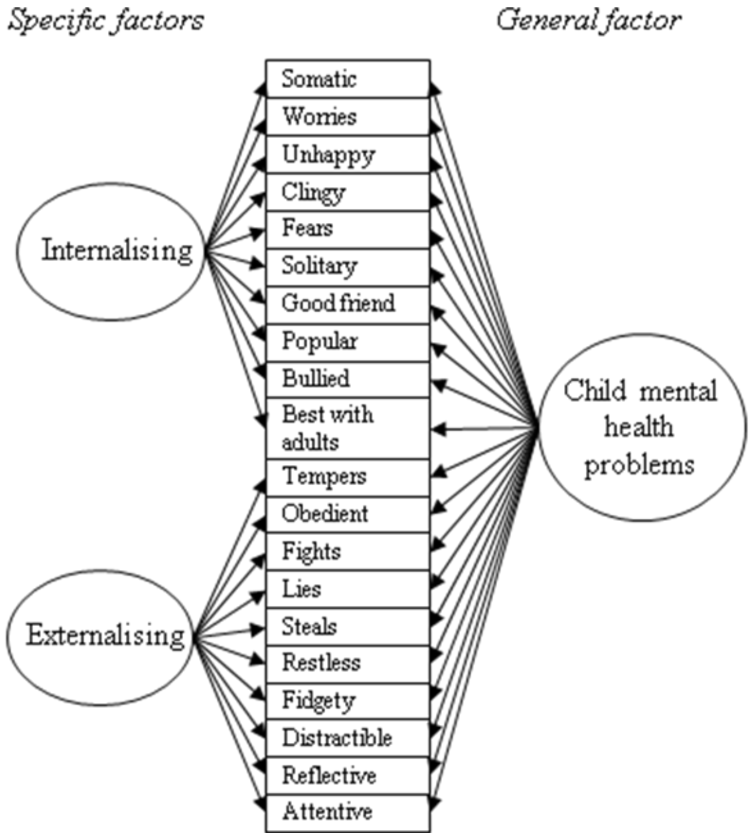
Both these EFAs and subsequent CFAs include all parents, teachers and children with full SDQ subscale data, and deal with the small number of missing items using MPlus's default pairwise present estimation.

Confirmatory factor analyses

Yet while I felt the EFAs provided useful preliminary analyses, EFAs cannot explicitly test the hypothesis that the SDQ factor structure is the same in Indians and Whites. For this purpose, I used a multi-group CFA. This is a method of assessing whether a particular factor structure relates in the same manner to the latent traits present in several populations [497]. If so then the factor structure is said to show measurement invariance which in this context implies strict between-group factorial invariance – that is, that all items have equivalent measurement parameters (thresholds, factor loadings and standard errors) across all groups. As in Section 8.1.2, I performed the CFA in MPlus5, using a multivariate probit analysis with the appropriate extension to ordinal data and using the default pairwise present approach to deal with the small amounts of missing data. For an example of the MPlus syntax I used for these multi-group CFAs, see Section 14.5.4 Appendix 2.

Section 8.1.3 reported that a two-factor (internalising/externalising) general-specific model provided acceptable fit to the parent, teacher and child SDQs in the total B-CAMHS sample (Table 8.3), with almost all items loading at >0.4. These results were based upon the entire B-CAMHS sample, but were almost identical after restricting the sample to only Indians and Whites (i.e. removing the other ethnic groups). I therefore took this model, illustrated in Figure 8.6, as my starting point when comparing the SDQ's factor structure between Indians and Whites.

Figure 8.6: Two-factor general-specific model[22]



Sample size and factor analyses

A limitation of both the EFAs and the multi-group CFA is the relatively small number of Indians: 389 parents, 306 teachers and 184 children. As discussed more fully in Appendix 1 p.375, this sample size is likely to be adequate for the parents and the teachers, but the child sample may be somewhat small. As inadequate sample sizes can lead to instability in the estimates of both EFA and CFAs, the results for the child SDQ should therefore be treated with some caution.

Construct validity of the SDQ

Construct validity of the SDQ subscales across informants

I performed multitrait-multimethod (MTMM) analyses to compare the construct validity of the internalising and externalising SDQ subscales in Indians and Whites. As when performing MTMM analyses in the full sample, I used Spearman's correlations and based each correlation upon all children with data for the SDQ(s) in question.

[22] For corresponding MPlus syntax, see Section 14.5.4 Appendix 2

Construct validity of the SDQ subscales relative to the DAWBA bands

In Section 8.1.3, I demonstrated in the full B-CAMHS sample that the SDQ subscales showed convergent and discriminant validity when predicting to DAWBA diagnosis at follow-up. There was not enough power to repeat these analyses separately in Indians, as only 142 Indian children were included in the follow-up surveys of whom three were diagnosed with emotional disorder, two with behavioural disorder and two with hyperactivity. Even among the 419 children included at baseline, there were only nine diagnoses of emotional disorder, five of behavioural disorder and one of hyperactivity.

I therefore modified the analysis strategy by changing the outcome to be the informant-specific emotional, behavioural and hyperactivity DAWBA bands measured at baseline. Each DAWBA band was used in turn as an outcome with the internalising and externalising SDQ subscales from the same informant as explanatory variables. For example, when the parent emotional DAWBA band was the outcome I entered the parent internalising and externalising subscales as explanatory variables. These DAWBA and SDQ measures are not independent because the SDQ subscales form part of the skip rules used when administering the DAWBA (see Chapter 5, p.124). Nevertheless, because this circularity applies equally to Indians and Whites, it does not undermine the intention of this analysis to *compare* the SDQ's construct validity between the two groups.

I ran these analyses separately in Indians and Whites for the parent, teacher and child SDQ. I excluded the small number of children (2-6%) with SDQ data but with missing data for one or more DAWBA bands from the same informant. This explains the slightly smaller sample sizes than in other SDQ analyses presented in this Section.

Association between symptoms and impact

The SDQ subscales consist of mental health symptoms, but the SDQ supplement also measures impact (see Chapter 5 p.126 and Section 14.2.1 Appendix 2). I calculated Spearman's coefficients for the association between the internalising/externalising subscales and the child's impact score. I then repeated this replacing the impact score with the single four-point item about burden to others.

Initially I performed these analyses for all Indian and all White children. I was, however, concerned that comparing these full Indian and White samples would be misleading because of the expectation that correlation coefficients will be lower in groups with lower variance [496]. Advantaged groups are expected to have lower variance on positively skewed scales like the SDQ because their scores are 'compressed' at the low end of the scale. As reported in Chapter 7, Indians are advantaged for the externalising subscales and their scores would therefore be expected to show lower variance than Whites. This was in fact observed: the standard deviation of the externalising subscale was 3.1 in Indians vs. 3.8 in Whites for the parent SDQ; 2.7 vs. 3.7 for the teacher SDQ; and 4.9 vs. 6.0 for the child SDQ.

For this reason, one would *a priori* expect lower correlations between symptoms and impact/burden in Indians than in Whites, even if the SDQ functioned equally well in both groups as a measure of mental health problems. I therefore decided additionally to present results on a group of White children who had a mental health advantage similar to the Indians. To do this, I selected at random a group of White children who were frequency-matched to the Indians by ±2 SDQ points for their internalising and externalising SDQ scores. I performed independent frequency matches for the parent, teacher and child SDQs. The resulting matched White samples had very similar subscale scores to the Indians, with the means differing by under 0.1 SDQ points for both the internalising and externalising subscales of all three informants.[23]

I therefore present correlation coefficients for symptoms and impact/burden for three groups: 1) the full sample of Whites, 2) a subsample of Whites frequency-matched to Indians for the internalising and externalising scores of the relevant informant and 3) the full sample of Indians. I then used Fisher's z-transformation (see Appendix 1, p.360) to compare the Spearman's coefficient of the Indians and the frequency-matched Whites.

[23] Achieving frequency matching meant reducing the White sample size by approximately two thirds, to around 5000 parents, 4000 teachers, and 2000 children. A more efficient approach would have been to assign each White child a probability weight based on the frequency of his or her SDQ scores in the Indians. Unfortunately this was not possible because Stata does not support probability weights when calculating Spearman's correlation coefficients. The White sample sizes remain very large even after frequency matching, however, meaning that in practice the reduced White sample size is unlikely to have any important implications for power.

8.2.3 Results

Internal factor structure of the SDQ in Indians and Whites

The factors derived empirically from two-factor EFAs in White and Indian children are presented in Table 14.20 (Section 14.5.2 Appendix 2) and their correspondence to the hypothesised internalising/externalising factors is presented in Table 8.8. As Table 8.8 shows, all items in the parent and teacher SDQs loaded on the expected subscale in Whites and all but one in Indians. There was only one cross-loading in parents (in Indians) and only three in teachers (two in Whites, one in Indians). This apparent similarity between Indians and Whites in their SDQ factor structures was then confirmed in the multi-group CFAs. As shown in Table 8.9, the two-factor general-specific model provided good or acceptable fit to the observed data in both parents and teachers (CFI and TLI>0.95, RMSEA<0.06). This provides evidence of measurement invariance with respect to ethnicity – i.e. that the loadings, thresholds and corresponding errors of each item in the SDQ are invariant across Indians and Whites, and furthermore that these symptoms correspond to latent traits which match the hypothesised internalising and externalising constructs.

In the child SDQ the correspondence of the empirically-derived factors with the hypothesised subscales was somewhat poorer in Indians, but there was again evidence of measurement invariance (CFI=0.959, TLI=0.947, RMSEA=0.046). The few unexpected loadings in the Indian child EFA may therefore simply reflect instability due to the somewhat small sample size.

Table 8.8: Correspondence of the empirically derived EFA factors with the hypothesised internalising/externalising constructs in Indians and Whites

			No. items loading correctly at >0.3		No. unexpected loadings at >0.3	
		N	**Internalising items**	**Externalising items**	**Internalising items**	**Externalising items**
Whites	**Parent**	16401	10/10	10/10	0/10	0/10
	Teacher	12865	10/10	10/10	2/10	0/10
	Child	6872	8/10	10/10	0/10	0/10
Indians	**Parent**	389	10/10	9/10	0/10	1/10
	Teacher	306	10/10	10/10	1/10	0/10
	Child	184	8/10	9/10	0/10	3/20

Table 8.9: Multi-group CFAs assessing measurement invariance between Indians and Whites

Informant	N	Factors	Model	CFI	TLI	RMSEA
Parent	16401 Whites, 389 Indian	Internalising/ externalising	Multi-group, general-specific	0.964	0.956	0.052
Teacher	12865 Whites, 306 Indians	Internalising/ externalising	Multi-group, general-specific	0.987	0.984	0.056
Child	6872 Whites, 184 Indians	Internalising/ externalising	Multi-group, general-specific	0.959	0.947	0.046

Construct validity of the SDQ

Construct validity of the SDQ across informants

Table 8.10 presents the MTMM analyses for Whites and Indians. In both groups the validity coefficients (grey) are larger than other coefficients in the same heterotrait block, as illustrated by the fact that the diamonds in Figure 8.7 are taller than they are wide. This indicates good convergent and discriminant validity within both groups, which is the main motivation for these analyses.

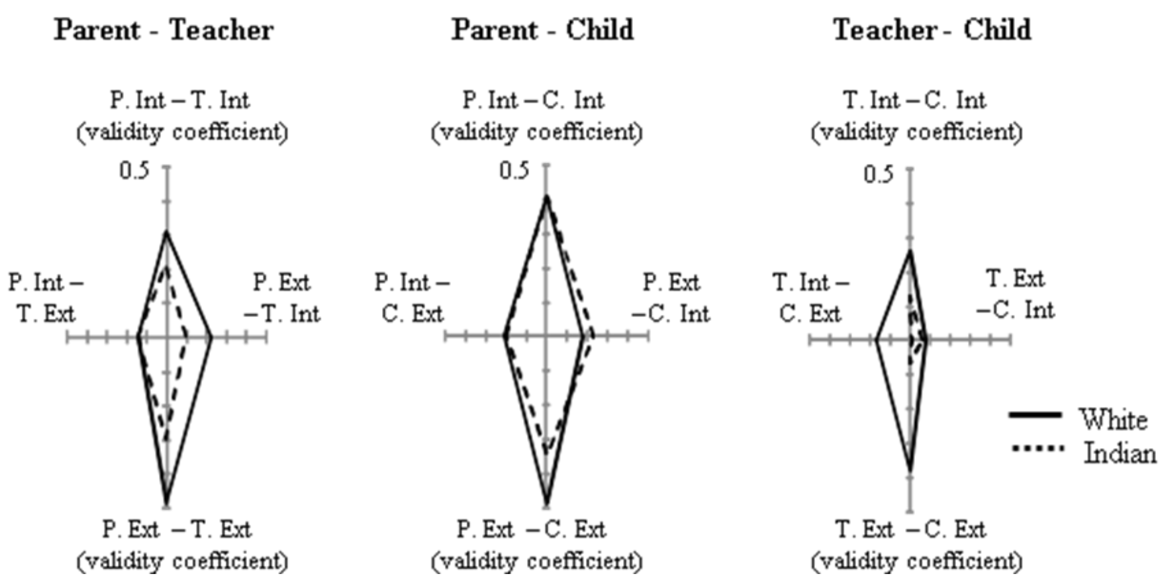
It is also of some interest to compare the absolute level of inter-informant correlation between Indians and Whites although, as with many analyses in this Chapter, any observed differences could have both methodological explanations and substantive explanations. Yet while Table 8.10 seems to suggest a lower inter-informant agreement in Indians than in Whites, making a 'fair' comparison between the groups is not straightforward. This is because of the same issue which affected the comparison of externalising scores and burden/impact above, namely the expectation that correlation coefficients will be lower in groups with lower variance [496]. In Section 14.5.4 Appendix 2 I present additional analyses which conclude that after addressing this issue there is no convincing evidence that the absolute values of the correlation coefficients are lower in Indians than Whites.

Table 8.10: MTMM analyses for Whites and Indians for the internalising and externalising subscales

			Parent		Teacher		Child	
			Int	Ext	Int	Ext	Int	Ext
Whites	**Parent**	**Int**						
		Ext	0.37					
	Teacher	**Int**	0.31	0.22				
		Ext	0.14	0.49	0.37			
	Child	**Int**	0.41	0.18	0.26	0.08		
		Ext	0.21	0.49	0.17	0.38	0.38	
Indians	**Parent**	**Int**						
		Ext	0.40					
	Teacher	**Int**	0.21	0.10				
		Ext	0.14	0.30	0.26			
	Child	**Int**	0.41	0.23	0.13	0.01		
		Ext	0.20	0.35	-0.06	0.07	0.36	

Int=internalising, Ext=externalising. Heterotrait blocks are outlined in bold, validity coefficients are shaded grey. For Whites, N=16401 parents; N=12865 teachers and N=6872 children. N=12839 for the parent-teacher comparison, N=6842 for the parent-child comparison and N=5199 for the teacher-child comparison. For Indians N=389 parents; N=306 teachers and N=184 children. N= 284 for the parent-teacher comparison, N=169 for the parent-child comparison and N=126 for the teacher-child comparison.

Figure 8.7: Radar plots comparing the construct validity of the internalising and externalising SDQ subscales in Whites and Indians



Each radar plot represents one heterotrait block from Table 8.10. P.=parent, T.=teacher, C.=child, Int=internalising subscale, Ext=externalising subscale

Construct validity of the SDQ subscales relative to the DAWBA bands

Table 8.11 shows which parent, teacher and child SDQ subscales independently predicted the DAWBA bands for that same informant. In many cases there was strong evidence (p<0.001) in Whites that the assumption of proportional odds was not met, but rather that the OR differed for different levels of the DAWBA band. The absolute differences in OR were not great, however, and the substantive conclusions regarding which subscale (internalising vs. externalising) was most strongly predictive did not change in non-proportional odds models. For simplicity, therefore, I present the proportional odds models in Table 8.11 and the equivalent non-proportional models in Table 14.25 (Section 14.5.6 Appendix 2).

In 8/8 cases in both Indians and Whites the expected SDQ subscale showed the strongest association with the DAWBA band by a substantial margin, with large OR (1.21 to 1.86 per SDQ point) which were highly significant (p<0.005 and usually p<0.001). Moreover, the absolute values of these OR were similar between Indians and Whites, with differences of 0.02 to 0.17 in the point estimates and with all confidence intervals overlapping. This therefore indicates comparable construct validity between Indians and Whites for the SDQ relative to the DAWBA bands. In other words, Indian and White informants were similar in the extent to which reporting more symptoms on the SDQ predicted also reporting more symptoms and impact on the far more detailed DAWBA.

Table 8.11: Internalising and externalising SDQ subscales as independent predictors of emotional, behavioural and hyperactivity DAWBA bands in Indians and Whites: proportional OR & 95%CI

		Informant-specific emotional DAWBA band		**Informant-specific behavioural DAWBA band**		**Informant-specific hyperactivity DAWBA band**	
		Whites	**Indians**	**Whites**	**Indians**	**Whites**	**Indians**
Parent	**Internalising SDQ**	1.50 (1.48, 1.52)***	1.37 (1.21, 1.55)***	1.03 (1.02, 1.05)***	0.94 (0.87, 1.01)	1.13 (1.10, 1.15)***	1.07 (0.93, 1.24)
	Externalising SDQ	1.03 (1.02, 1.04)***	1.07 (0.93, 1.24)	1.48 (1.47, 1.50)***	1.38 (1.28, 1.49)***	1.65 (1.63, 1.68)***	1.64 (1.40, 1.92)***
Teacher	**Internalising SDQ**	1.50 (1.47, 1.53)***	1.62 (1.40, 1.88)***	1.12 (1.10, 1.14)***	1.02 (0.93, 1.12)	1.09 (1.07, 1.11)***	1.08 (0.97, 1.21)
	Externalising SDQ	1.18 (1.16, 1.20)***	1.42 (1.24, 1.62)***	1.60 (1.58, 1.63)***	1.62 (1.44, 1.82)***	1.81 (1.77, 1.84)***	1.86 (1.62, 2.13)***
Child	**Internalising SDQ**	1.40 (1.37, 1.43)***	1.43 (1.27, 1.6)***	0.98 (0.96, 1.00)*	1.09 (0.97, 1.22)***	–	–
	Externalising SDQ	1.10 (1.08, 1.12)***	1.17 (0.98, 1.39)***	1.38 (1.35, 1.40)***	1.21 (1.05, 1.37)**	–	–

*p<0.05, **p<0.01, ***p<0.001. Subscales shaded grey are expected *a priori* to be the strongest predictors. N=16 338 Whites and 365 Indians for the parent DAWBA bands, 12 312 Whites and 293 Indians for the teacher DAWBA bands; and 6781 Whites and 181 Indians for the child DAWBA bands. Note no hyperactivity DAWBA band exists for children. See Table 14.25 (Section 14.5.6, Appendix 2) for the equivalent non-proportional odds model.

Association between symptoms and impact

Table 8.12 presents the association of the internalising and externalising subscales with the impact and burden reported on the SDQ. All correlation coefficients are positive and of moderate size (0.3-0.6) indicating the substantial negative effect of both internalising and externalising symptoms upon children's lives and the lives of those who care for them. These correlation coefficients are generally similar between Indians and the frequency-matched White sample, with no evidence (p<0.05) of a difference in 11/12 comparisons, and only weak evidence in the twelfth (p=0.03). This therefore provides evidence that the SDQ subscales measure emic mental health problems with important implications for child well-being, and that this is true in both Indians and Whites.

Table 8.12: Association between the internalising/externalising subscales with impact and burden

	Parent			**Teacher**			**Child**		
	White (full)	**White (match)**	**Indian**	**White (full)**	**White (match)**	**Indian**	**White (full)**	**White (match)**	**Indian**
N	16393	4991	385	12744	3911	302	6812	2199	180
Impact-internalising	0.42	0.43	0.36	0.44	0.44	0.39	0.31	0.32	0.31
Impact-externalising	0.43	0.35	0.31	0.56	0.48	0.44	0.34	0.31	0.34
Burden-internalising	0.43	0.44	0.34*	0.40	0.40	0.32	0.31	0.32	0.31
Burden-externalising	0.47	0.40	0.33	0.62	0.56	0.54	0.38	0.37	0.32

*p<0.05, **p<0.01, ***p<0.001, testing for equality between Indians and the matched White sample.
Full=full sample, match=frequency-matched sample for internalising and externalising scores, matched using the SDQs of the informant in question.

8.2.4 Discussion and conclusions: the common mental health problems are comparable between Whites and Indians

As summarised in Section 5.3.2 Chapter 5, the SDQ has been used with apparent success in the Indian subcontinent but rigorous psychometric assessments have typically been lacking. Likewise, as highlighted in Chapter 4 (p.106), one of the key failings of the existing mental health literature comparing different ethnic groups in Britain is the near universal failure to validate the mental health measures in each of the groups under consideration.

This Section addresses this issue for the ethnic groups which form the focus of this PhD. It demonstrates that the SDQ's factor structure shows measurement invariance between Indians and Whites. The Indian and White SDQs also showed comparable construct validity across different informants, relative to the DAWBA bands, and relative to reported impact and burden. Together, these results provide strong evidence that Indians and Whites experience similar child mental health problems.

Had the results of these analyses been otherwise, this could have implied a violation of the universalist assumption with regards to internalising and externalising problems in Indians and Whites in B-CAMHS. Alternatively it could simply have implied that the SDQ was not a valid measure of internalising and externalising problems in one or both populations. Distinguishing between these two alternatives might not have been possible. As it is, however, the findings in this Section imply *both* that internalising and externalising problems do exist in Indians and Whites *and* that these can be measured using the SDQ.

This conclusion is of central importance in justifying further comparisons of internalising and externalising mental health problems Indians and Whites. Nonetheless, the comparison of Indians and Whites could still be undermined by selection or information biases. I have previously demonstrated in Section 6.3, Chapter 6 that there was no evidence of selection biases between Indians and Whites which might explain their mental health advantage. I now turn to the issue of information biases in the final section of this Chapter.

8.3 Potential biases in the mental health assessments of Indians and Whites

8.3.1 Deviations from criterion equivalence (threshold bias)

One form of information bias which can jeopardise inter-group comparisons is a deviation from criterion equivalence – that is, a systematic difference between groups in the thresholds for endorsing items. Such systematic differences could apply only to particular items (an item bias) or could apply to most or all items (a method bias).

The demonstration of measurement invariance between Indians and Whites in Section 8.2 (Table 8.9) provides evidence against item biases. Evidence against a method bias comes from the fact that all informants report the Indian advantage (see Chapter 7). Crucially, this includes teachers, the culturally 'external' informant. This indicates that the Indian advantage cannot simply be attributed to systematic differences between the norms of Indian and White parents.

Further important evidence against a method bias comes from the fact that the Indian advantage is observed not only for the SDQ but also for the DAWBA bands and DAWBA diagnoses. Moreover, as reported in Section 8.2 (Table 8.11), the relationship between number of SDQ symptoms and number of same-informant DAWBA symptoms is very similar in Indians and Whites. This is important because, as argued in Chapter 5 (p.135), the longer and more detailed DAWBA may be less prone to cross-cultural bias than the SDQ. Certainly this seemed to apply to a comparison of British and Norwegian children which my colleagues and I conducted. As we showed, Norwegian informants systematically underreported emotional problems on the SDQ, but this reporting bias disappeared when using the DAWBA bands ([230], Appendix 3).[24] The fact that the Indian

[24] The motivation for this comparison was to investigate whether a large apparent mental health advantage in one group (Norwegian children) relative to another (British children) reflected a real difference or a reporting bias. It was therefore conceptually very similar to the aim of this Chapter. The Norwegian sample was larger and more informative, however, including full measures on 1024 children after oversampling for mental health problems. It was therefore possible to examine reporting biases in the SDQ questionnaires when judged against emotional, behavioural or hyperactivity DAWBA diagnosis, and against parental reports of

SDQ advantage persists on the DAWBA therefore provides evidence against a similar explanation underlying the Indian advantage.

8.3.2 Other possible forms of systematic bias: motivating questions for this section

I therefore believe that the apparent Indian advantage cannot be attributed to a deviation from criterion equivalence. I now examine three other potential cross-cultural biases:

1. Disclosure to open-ended questions.
2. Parent informant type.
3. Parent language of interview.

As discussed in Section 2.4 Chapter 2, the first is a *mechanism* of information bias, and I examine this issue for all informants. The second and third factors are plausible *sources* of information bias which apply specifically to the responses of Indian and White parents.

8.3.3 Methods

Disclosure in the open-ended DAWBA transcripts

Matching Indian and White children

I matched the 419 Indian children to 419 randomly-selected White children, matching for:

- Gender.
- Age (±2 years).
- Survey year (1999 vs. 2004).
- 'Any disorder' parent DAWBA band.
- 'Any disorder' teacher DAWBA band.
- 'Any disorder' child DAWBA band.
- Whether the child received a DAWBA diagnosis. In 11/14 cases it was possible to match this on disorder type (emotional, behavioural, hyperactivity or autistic disorder).

---

definite or severe problems. By contrast, these outcomes were too rare to be used in the smaller (N=419) and non-selected group of Indians in B-CAMHS.

I matched for these factors to prevent them confounding the comparison between Indians and Whites of disclosure in the open-ended DAWBA transcripts. Matching for mental health status was particularly crucial because this is a central determinant of how much informants report in the transcript. The informant-specific DAWBA bands seemed the most appropriate mental health matching criteria because they provide a direct measure of symptoms reported by that same informant in the DAWBA interview. Matching in this way therefore allows investigation of whether, given a similar level of reported concern in the structured DAWBA interview, Indian and White informants provided the same amount of narrative information when asked to describe any problems in more detail.

Extracting data on disclosure in transcripts

For each child, I extracted those sections of the DAWBA transcripts in which informants were asked to describe any areas of difficulty. I then used the MSWord 'word count' function to quantify how much they said. I independently repeated extracting number of words for each informant for a randomly selected 100 children. Re-test reliability was excellent, with a Spearman's correlation of 1.00 for parents and teachers, and 0.97 for children.

Word count is clearly a crude metric of information, and I additionally explored using number of sentences. I did not pursue this approach, however, because the opinion of two child psychiatrists who had rated the B-CAMHS DAWBAs was that word count was a more helpful proxy for amount of information. In addition, deciding where sentences end in free text transcripts is not straightforward.

Comparing transcript length between Indians and Whites

The distributions of number of words reported by parents, teachers and children were all highly positively skewed, with most informants saying nothing or very little but a few giving very long accounts. I therefore compared the raw number of words between Indians and Whites using a non-parametric Wilcoxon matched pairs signed-ranks test.

The expectation in both ethnic groups is for reported concern in the structured DAWBA to be positively associated with disclosure in the transcripts. If Indians and Whites differed in this respect then this could suggest non-equivalence in the nature and/or utility of the transcript information. I therefore calculated Spearman’s coefficients for the agreement between DAWBA band and transcript length in both ethnic groups, and compared these using Fisher's z-transformation.

Parent informant

Type of informant

Both B-CAMHS surveys collected information on how the ‘parent’ informant was related to the child (mother, father etc). I used this information to compare the distribution of parent informant types between Indians and Whites.

Mental health by informant

A difference between Indians and Whites in the distribution of parent informant types would carry the *potential* to bias inter-group comparisons. It would, however, not in itself be enough actually to cause a bias. Rather, there would also need to be a reporting bias between different parent informants (e.g. fathers systematically over-report symptoms relative to mothers). Identifying such biases is complicated by the existence of plausible substantive reasons for differences between the mental health assessments of mother informants vs. father informants. For example, children with father informants might disproportionately come from socio-economically disadvantaged families in which the father was unemployed. In this case parent informant type would not be a source of bias (i.e. creating spurious differences between Indians and Whites) but rather would be a marker for confounders which might explain real ethnic differences.

To investigate this issue, I calculated the prevalence or mean of a range of mental health outcomes after stratifying by parent informant type. The outcomes were DAWBA diagnosis of internalising (emotional) disorder; DAWBA diagnosis of externalising (behavioural or hyperactive) disorder; and the internalising and externalising subscales of the parent, teacher and child SDQs. I included the teacher and child SDQs because these external informants should not be affected by reporting biases between parent informant types. Any differences observed by teacher/child-report as well as by parent-report would therefore suggest a ‘confounding’ rather than a ‘bias’ explanation.

I present these DAWBA and SDQ outcomes for the combined sample of Indians and Whites, and also for Indians and Whites separately. In the combined Indian and White sample I fitted logistic/linear regression models to investigate whether these mental health outcomes were predicted by informant type. I also thought it plausible that the reasons for (say) fathers vs. mothers to be the respondent might differ between Indians and Whites, and therefore that the effect of parent informant type upon mental health assessment might differ between the groups. I therefore tested for interactions between parent informant type and Indian ethnicity in each regression model, using a 1% significance cut-off to reduce the probability of chance findings through multiple testing.

Language of interview

Proportion of translated parent SDQs

Both B-CAMHS surveys recorded whether the parent completed the SDQ in a language other than English and I used this information to compare Indians and Whites. I also calculated what proportion of those parents who completed the SDQ in translation went on to complete the parent DAWBA, as a means of assessing whether language seemed to have been a major barrier to their participation.

Mental health by language of response

As for parent informant type, a difference in the proportion of SDQs completed in translation would only bias inter-ethnic comparisons if there were also a reporting bias between English and non-English SDQs. Ideally, one would use a method similar to that described above for parent informant type to investigate each non-English language individually (as the SDQ might be biased for some languages but not others). This was not possible, however, because there was no overlap in the non-English languages chosen by Indian and White parents. Moreover, in fact very few parents completed SDQs in translation. The resulting power limitation was exacerbated by ONS having no record of the scores for the translated SDQs administered in B-CAMHS99. All that proved possible was therefore a comparison *within* Indians of the mental health of children whose parents completed the SDQ in English vs. those who completed it in translation.

8.3.4 Results

Disclosure in the open-ended DAWBA transcripts

The distribution of word counts was strikingly similar between Indians and Whites, with no evidence of a difference for any informant (Table 8.13). Moreover, the expected positive correlation between DAWBA band and transcript length was always observed, and there was never evidence that the strength of this association differed between Indians and Whites (Table 8.14). There was therefore no evidence that, given a similar level of concern reported in the structured DAWBA sections, Indian and White informants differed in how much information they disclosed in the DAWBA transcripts.

Table 8.13: Transcript length for matched Indian and White samples

	Parent transcript		Teacher transcript		Child transcript	
	White	**Indian**	**White**	**Indian**	**White**	**Indian**
N	370	370	311	311	185	185
No. transcripts containing:						
0 words [transcript blank]	291	282	237	246	135	125
1-49 words	28	22	58	54	15	23
50-199 words	37	52	15	11	31	31
≥200 words	14	14	1	0	4	6
P-value for difference (Wilcoxon matched pairs signed-ranks test)	*0.25*		*0.18*		*0.21*	

Table 8.14: Correlation between word count and 'any disorder' DAWBA band in matched Indian and White samples

	Parent transcript		Teacher transcript		Child transcript	
	White	**Indian**	**White**	**Indian**	**White**	**Indian**
N	370	370	311	311	185	185
Spearman's coefficient	0.36	0.37	0.21	0.12	0.34	0.46
P-value for difference(Fisher's z-transformation)	*0.88*		*0.25*		*0.17*	

Parent informant

Type of informant

There was strong evidence of a difference between the informant types of Whites and Indians (p<0.001). As shown in Table 8.15, this resulted from fathers making up a far higher proportion of parent respondents among Indians (18.1%) than Whites (4.3%). By contrast, both ethnic groups contained very few stepparents, grandparents and 'other' informant types. I therefore excluded these informant types from subsequent analyses, as being too rare to allow meaningful comparisons.

Table 8.15: Parent informant types for Whites and Indians

	Full population of Whites and Indians N (%)	**White N (%)**	**Indian N (%)**
Biological mother	15 871 (94.1%)	15 532 (94.4%)	339 (81.2%)
Biological father	775 (4.6%)	699 (4.3%)	76 (18.1%)
Stepparent	115 (0.7%)	114 (0.7%)	1 (0.2%)
Grandparent	55 (0.3%)	54 (0.3%)	1 (0.2%)
Other	51 (0.3%)	50 (0.3%)	1 (0.2%)
Total	16 867(100%)	16 449 (100%)	418 (100%)

Parent informant type missing for one Indian parent.

Mental health by informant (mothers vs. fathers)

The proportion of mothers and fathers was very similar for children with an internalising/externalising disorder compared to children without (Table 8.16). This suggests that parent informant type does not predict child disorder status, a conclusion borne out by formal statistical testing. There was no evidence (p>0.1) of a difference between children with mother informants vs. children with father informants in the prevalence of internalising DAWBA diagnoses, externalising DAWBA diagnoses, parent SDQ scores or child SDQ scores (Table 8.17). The only evidence of a difference between parent informants came from the teacher SDQ, which indicated more externalising and perhaps more internalising problems in children with father informants compared to children with mother informants. As teachers are an external informant this is unlikely to reflect a reporting bias. Nevertheless, the fact that the effect is not replicated on other measures (including the child SDQ, the other external informant) suggests it is also unlikely to be a marker for a genuine confounder. It therefore seems most likely to represent a chance finding.

In no case was there evidence at the 1% level of an interaction between Indian ethnicity and parent informant type. Rather, as shown in the third and fourth columns of Table 8.17, Whites and Indians both showed little difference by informant type for most mental health outcomes. It is therefore unsurprising that stratifying by parent informant type had very little effect on the Indian mental health advantage for externalising disorders/problems. Rather within both mother and father informants, the prevalence of externalising disorder was substantially lower in Indians than Whites. So too were all SDQ scores. These data therefore provide no suggestion that the higher proportion of father informants among Indians is the explanation for the observed Indian advantage.

Table 8.16: Proportion of mothers and fathers, stratified by disorder status for internalising/externalising DAWBA diagnosis

		Full population N (%)		**White N (%)**		**Indian N (%)**	
		No disorder	**Disorder**	**No disorder**	**Disorder**	**No disorder**	**Disorder**
Internalising disorder	**Mother**	15232 (95.3%)	639 (96.2%)	14900 (95.6%)	632 (96.5%)	332 (81.8%)	7 (77.8%)
	Father	750 (4.7%)	25 (3.8%)	676 (4.3%)	23 (3.5%)	74 (18.2%)	2 (22.2%)
	Total	15982 (100%)	664 (100%)	15776 (100%)	655 (100%)	406 (100%)	9 (100%)
Externalising disorder	**Mother**	14926 (95.4%)	945 (95.1%)	14593 (95.7%)	939 (95.0%)	333 (81.4%)	6 (100%)
	Father	726 (4.6%)	49 (4.9%)	650 (4.3%)	49 (5.0%)	76 (18.6%)	0 (0%)
	Total	15652 (100%)	994 (100%)	15243 (100%)	988 (100%)	409 (100%)	6 (100%)

Table 8.17: Prevalence of DAWBA diagnosis and mean parent, teacher and child SDQ scores by parent informant-type

	Informant	Full population prevalence (%)/mean	P-value for association with parent informant type in full population†	White prevalence (%)/mean	Indian prevalence (%)/mean
Prevalence DAWBA internalising diagnosis	**Mother** (15517 White, 339 Indian)	4.0 (3.7, 4.4)	0.26	4.1 (3.7, 4.4)	2.1 (1.0, 4.4)
	Father (669 White, 76 Indian)	3.2 (2.2, 4.7)		3.3 (2.2, 4.9)	2.6 (0.8, 8.7)
Prevalence DAWBA externalising diagnosis	**Mother** (15517 White, 339 Indian)	6.0 (5.6, 6.4)	0.68	6.0 (5.7, 6.5)	1.8 (0.7, 4.2)
	Father (669 White, 76 Indian)	6.3 (4.8, 8.4)		7.0 (5.3, 9.3)	[zero cases]
Mean parent internalising score	**Mother** (15476 White, 318 Indian)	3.32 (3.26, 3.38)	0.56	3.31 (3.25, 3.37)	3.69 (3.22, 4.16)
	Father (694 White, 69 Indian)	3.26 (3.06, 3.45)		3.24 (3.04, 3.45)	3.36 (2.64, 4.09)
Mean parent externalising score	**Mother** (15476 White, 318 Indian)	4.89 (4.82, 4.97)	0.67	4.91 (4.83, 4.99)	3.95 (3.59, 4.31)
	Father (694 White, 69 Indian)	4.95 (4.68, 5.22)		5.05 (4.77, 5.34)	3.90 (3.21, 4.59)
Mean teacher internalising score	**Mother** (12112 White, 244 Indian)	2.79 (2.73, 2.85)	0.02	2.79 (2.73, 2.85)	2.58 (2.05, 3.11)
	Father (511White, 54 Indian)	3.14 (2.85, 3.43)		3.25 (2.93, 3.56)	2.13 (1.50, 2.76)
Mean teacher externalising score	**Mother** (12112 White, 244 Indian)	3.62 (3.54, 3.69)	0.003	3.64 (3.56, 3.71)	2.52 (2.10, 2.95)
	Father (511White, 54 Indian)	4.18 (3.81, 4.56)		4.27 (3.87, 4.68)	3.33 (2.38, 4.29)
Mean child internalising score	**Mother** (6422 White, 150 Indian)	4.20 (4.12, 4.27)	0.12	4.20 (4.12, 4.28)	4.00 (3.56, 4.44)
	Father (314 White, 30 Indian)	4.45 (4.14, 4.76)		4.43 (4.10, 4.75)	4.70 (3.71, 5.69)
Mean child externalising score	**Mother** (6422 White, 150 Indian)	5.98 (5.90, 6.07)	0.43	6.01 (5.92, 6.09)	5.02 (4.49, 5.55)
	Father (314 White, 30 Indian)	6.13 (5.76, 6.50)		6.32 (5.94, 6.70)	4.13 (3.17, 5.10)

†Logistic regression was used for the DAWBA diagnoses, linear for the SDQ subscales. In no case was there evidence at the 1% level of an interaction between Indian ethnicity and parent informant type

Language of interview

Proportion of translated parent SDQs

A substantially higher proportion of Indian parents than White parents completed SDQs in languages other than English (10.0% of Indian parents vs. 0.1% in White, $p<0.001$; see Table 8.18). Among Indians the most common non-English SDQ languages were Gujarati (N=26) and Punjabi (N=11) while among Whites no language was chosen by more than three parents.

Table 8.18: Use of translated SDQs for White and Indian parents

	White N (%)	**Indian N (%)**
English SDQ	16 430 (99.9%)	377 (90.0%)
Non-English SDQ	19 (0.1%)	42 (10.0%)
Total	16 449 (100%)	419 (100%)

None of the 19 White or 42 Indian parents who completed a translated SDQ went on to complete the parent DAWBA. This suggests language proved an important barrier to the collection of mental health data on these children. The DAWBA diagnoses assigned to these children were therefore based on teacher and child interviews. 23 further children with translated parent SDQs did not have teacher or child data and so do not feature as participants in B-CAMHS survey. According to the Onomap name matching program (see Section 6.3 Chapter 6), none of these non-participating children had White names and three had Indian names. It therefore seems that problems with English seem were a barrier to the participation of some parents, but that few children were excluded entirely for this reason.

Mental health by language of response

With only 42 Indian parents answering in non-English languages, there was very little power to detect differences in prevalence of disorder by parental language. The most that can be said is that the observation of one DAWBA diagnosis among these 42 children (2.4%) is consistent with the 13/377 (3.5%) prevalence among Indian children whose parents responded in English.

The dimensional SDQ offers somewhat more power to compare Indians by parent language of response, although unfortunately ONS only had a record of translated parent SDQ scores for B-CAMHS04. As Table 8.19 shows, there was strong evidence that Indian parents reported higher internalising scores on the translated SDQs (5.77 vs. 3.08, $p<0.001$), and also a trend in this direction for the externalising scores (4.72 vs. 3.34, p=0.09). By contrast, there was no evidence of any such differences on the teacher or child SDQs. This remained the case after restricting the analyses of the teacher and child SDQs to B-CAMHS04.

Table 8.19: Mental health by SDQ response language for Indian children

		N	Mean (95%CI)	P-value for difference†
Mean parent internalising score (B-CAMHS04)	**English SDQ**	168	3.08 (2.54, 3.62)	<0.001
	Translated SDQ	22	5.77 (4.60, 6.94)	
Mean parent externalising score (B-CAMHS04)	**English SDQ**	168	3.34 (2.97, 3.72)	0.09
	Translated SDQ	22	4.72 (3.24, 6.20)	
Mean teacher internalising score	**English SDQ**	267	2.55 (2.06, 3.04)	0.63
	Translated SDQ	35	2.31 (1.42, 3.21)	
Mean teacher externalising score	**English SDQ**	267	2.64 (2.28, 3.01)	0.88
	Translated SDQ	35	2.57 (1.63, 3.52)	
Mean child internalising score	**English SDQ**	163	4.12 (3.65, 4.58)	0.49
	Translated SDQ	19	4.44 (3.67, 5.21)	
Mean child externalising score	**English SDQ**	163	4.86 (4.38, 5.35)	0.78
	Translated SDQ	19	4.69 (3.52, 5.87)	

†P-value for difference between English and translated SDQ scores, calculated using a T-test. The substantive findings were unchanged when the p-values were recalculated after taking zero-skew logs.

That no differences were observed on the teacher and child SDQs suggests that reporting bias may explain the higher scores reported by parents completing non-English SDQs. This could reflect differences in the norms of Indian parents who answer in translation rather than in English or could simply reflect inadequate translation of the SDQ. The latter is certainly plausible given that several translated SDQs, including those in Gujarati and Punjabi, were professionally translated by ONS but never validated in detail (Robert Goodman – personal communication).

Does this reporting bias explain the apparent disadvantage of Indian girls for parent-reported internalising problems?

There was thus evidence of a reporting bias among Indians completing the parent SDQ in translation, and this was particularly strong for the internalising scale. This is especially interesting because in Chapter 7 the parent-reported internalising subscale was the only mental health outcome which showed any evidence of an Indian disadvantage. This apparent disadvantage, which appeared to apply particularly to girls, was not replicated in other informants or using the DAWBA measures. It was also not consistent with the findings of previous large studies.

To investigate whether this anomalous finding could be explained by the reporting bias identified above, I repeated the analyses presented in Table 7.4, Chapter 7 after excluding the parent SDQs in translation. As Table 8.20 shows, after excluding the translated SDQs, there was no longer evidence of an interaction with ethnicity at the 1% level (p=0.03 vs. p=0.007 previously in Chapter 7), and when boys and girls were analysed together there was no evidence of a difference between Indians and Whites. Moreover, even when stratifying by gender there was now only borderline evidence for a disadvantage in girls (p=0.06 vs. p=0.005 previously). This therefore suggests that the anomalous finding in Chapter 7 for parent-reported internalising problems was indeed driven by reporting bias from the parent SDQs in translation.

Table 8.20: Comparison of mean parent internalising SDQ scales, with and without SDQs in translation

	All children, as in Table 7.4, Chapter 7 (N=16386 for White, N=389 for Indian)			**Children with parent SDQs in English (N=16381 for White, N=367 for Indian)**		
	White mean	**Indian mean**	**Regression coefficient & 95%CI for White (vs. Indian) ethnicity**	**White mean**	**Indian mean**	**Regression coefficient & 95%CI for White (vs. Indian) ethnicity**
All children	3.31	3.63	*Gender interaction (p=0.007)*	3.31	3.51	-0.20 (-0.66, 0.25) [*gender interaction p=0.03]*
Boys	3.33	3.24	0.10 (-0.41, 0.61)	3.33	3.21	0.13 (-0.39, 0.65)
Girls	3.30	4.05	-0.77 (-1.31, -0.24)**	3.30	3.86	-0.57 (-1.15, 0.01) [p=0.06]

*p<0.05, **p<0.01, ***p<0.001. All analyses adjusted for age, gender and survey year.

8.3.5 Discussion and conclusions: bias cannot explain the Indian mental health advantage

To summarise, after matching for symptoms and impact reported in the fully-structured DAWBA sections, Indian and White informants were very similar in the amount they disclosed in the open-ended transcripts. There was therefore no evidence that lower rates of disclosure among Indian parents, teachers or children might have led to mental health problems being missed more frequently in Indian children. This is important because the observation of the Indian mental health advantage when using the ‘gold standard’ of DAWBA diagnosis represents strong evidence that the Indian advantage observed on the SDQ is real rather than a reporting bias. This finding is also useful in supporting the DAWBA’s face validity as a cross-cultural measurement tool.

Among the Indian and White parents participating in B-CAMHS, a substantially higher proportion of the Indian parents were fathers. There was little evidence, however, of any reporting bias between mothers and fathers. Rather the mental health of children whose fathers participated was, at a group level, very similar to that of children whose mothers participated. This was true as judged both by parent-report and by external assessments, and was also true in both Indians and Whites. It is therefore unsurprising that the mental health difference between Indians and Whites remained almost unchanged after stratifying by parent informant type. This indicates that the higher proportion of Indian father participants is not a plausible explanation for the observed Indian advantage. It is also of broader interest in implying that mothers and fathers in Indian and White families make similar assessments of the mental health of their children. This is consistent with the comparatively high correlation (0.55-0.65) between the mental health assessments of mother-father pairs reported in a previous study of 125 Indian and White children from South-East England [387].

By contrast, there was evidence that SDQ language introduced bias into the comparison of Indians and Whites. 10.0% of Indian parents completed the SDQ a non-English language, as compared to 0.1% of Whites. These figures are similar to the Millennium Cohort Survey's finding that 10.5% of Indian families and 0.5% of White families spoke only non-English languages at home [242]. My analyses of the effects of this difference were limited by small samples sizes. It was striking, however, that despite the low power there was evidence that Indian parents completing translated SDQs reported substantially higher scores than Indian parents completing SDQs in English. There was no evidence of such a difference by the external informants of teachers and children, suggesting the higher parent scores on the translated SDQs reflect a reporting bias. This provides 'proof of principle' of the fact that using multiple informants, including the external informants of teachers, can identify instances of reporting bias.

Given this apparent reporting bias, I have decided to exclude from all subsequent analyses the four White and 22 Indian parents who completed translated SDQs in B-CAMHS04 (those from B-CAMHS99 are *de facto* missing). Notably, revisiting the analyses reported in Chapter 7 in the light of this exclusion removed the one anomalous instance of an Indian disadvantage. Clearly in future surveys it will be important to assess the translation and psychometric properties of the non-English SDQs more carefully.

8.4 Chapter summary and conclusion

The analyses presented in this Chapter generally support the validity of using the DAWBA and SDQ to compare internalising and externalising problems between Indians and Whites. Internalising and externalising mental health problems appear to be meaningful entities in both ethnic groups, with very similar SDQ factor structures, construct validity and associations between symptoms and impact. This and the previous Chapter also indicate that deviations from criterion equivalence (threshold biases) are unlikely to explain the Indian advantage. This is evidenced by the fact that the Indian advantage is equally apparent by teacher report (the external informant) and in the more detailed DAWBA interview. There is also no evidence that systematic biases related to disclosure rates or parent informant type can explain the Indian mental health advantage. Instead the apparent reporting bias for the non-English parent SDQs was the only analysis in the whole Chapter which gave any grounds for concern. This is easily addressed by excluding the small number of translated SDQs involved.

Chapter 7 demonstrated consistent evidence across measures and across informants of a large Indian advantage for externalising problems, but little or no evidence of a difference for internalising problems. Taking this in conjunction with the results reported in this Chapter, I conclude that there is good evidence that Indian children in B-CAMHS are no different to Whites for internalising problems but really do enjoy a lower prevalence of externalising problems.

Chapter 9 Aim three, part one: Conceptual models and explanatory variables for understanding the Indian advantage

The previous two Chapters addressed the first two aims of my thesis by demonstrating that the B-CAMHS Indian mental health advantage is specific to externalising problems and is unlikely to result from ethnic differences in how mental health is conceptualised or reported. Yet as argued in Section 3.1.3 Chapter 3, the *description* of ethnic variation in health is not an *explanation* for that variation, but rather should be a starting point for further investigation into operative causal mechanisms. This brings me to this PhD's third aim, namely to investigate whether any real Indian advantage in B-CAMHS can be explained by the child, family, school and area characteristics of Indian children.

This Chapter outlines how I intend to investigate the causes of the Indian advantage for externalising problems. First I justify my choice of mental health outcome. I then present a conceptual model for how the child, family, school and area variables collected in B-CAMHS may be related to ethnicity, and describe these individual explanatory variables in detail. This provides a foundation for the subsequent two Chapters which seek to explain the Indian advantage through univariable and multivariable analyses.

9.1 Outcomes and conceptual models for substantive analyses

9.1.1 Mental health outcomes

In addressing the third aim of my PhD, I focus upon externalising problems because it is for these that Indians have a mental health advantage. My primary mental health outcome is the externalising subscale of the parent SDQ ('parent externalising score'), after excluding the small number of translated SDQs. This is available for almost all B-CAMHS participants and consistently showed good psychometric properties in both ethnic groups in Section 8.2 Chapter 8. I shall also use the teacher externalising score as a secondary outcome measure. This likewise showed good psychometric properties in both ethnic groups in Chapter 8, and I shall use it to assess how far findings from teachers replicate those from parents. This is valuable because teachers are arguably less likely to show cross cultural bias than parents. Teachers also provide a mental health assessment which is independent of most of the risk factor data, which was largely provided by parents.

A key advantage of the parent and teacher externalising scores is that they are dimensional measures. By contrast, despite being the single best measure in B-CAMHS of severe mental health problems, the DAWBA generated only six diagnoses of externalising disorder for Indians. It is therefore not well powered as an outcome for the kinds of analyses presented in Chapter 10 and Chapter 11. Nevertheless, as a final sensitivity analysis I will seek replicate my findings from the parent and teacher SDQs using DAWBA diagnoses and also the DAWBA bands and the child SDQ. I will also present final analyses using internalising disorders/problems as an outcome in order to verify that adjusting for child, family, school and area factors does not unmask an unexplained difference between Indians and Whites.

9.1.2 A hierarchical conceptual model for characteristics which might mediate the Indian advantage

I believe that explicitly hierarchical models are useful when investigating the effects of 'upstream' variables like ethnicity [498]. I therefore grouped the many B-CAMHS explanatory variables both by their substantive content and by how proximate a causal factor I hypothesised them to be.[25] The resulting conceptual model is presented in Table 9.1 and Figure 9.1. This model hypothesises three Levels at which factors which may influence a child's mental health:

- Level 1 includes three domains of explanatory variables: area characteristics, school characteristics and family socio-economic position (SEP). I see all three domains as relating to structural and/or cultural factors which may have a direct effect upon externalising problems or may be mediated by more proximate family and child factors. These domains are also all plausibly related to historical processes which

[25] By contrast, the only previous detailed multivariable analysis of B-CAMHS simply groups the variables into 'child', 'family', 'area' and 'school' factors 417. Ford, T., R. Goodman, and H. Meltzer, *The relative importance of child, family, school and neighbourhood correlates of childhood psychiatric disorder*. Soc Psychiatry Psychiatr Epidemiol, 2004. **39**(6): p. 487-96.. My analysis also differs from this previous analysis in containing substantially more explanatory variables.

have led to ethnic differences in geographic or socio-economic profiles (see Figure 9.1).

- Level 2 includes the two domains of family composition and family stress. My hypothesis is that the Level 2 variables may be partly caused by the Level 1 variables. I also hypothesise that the Level 2 variables may have direct effects on externalising problems but may alternatively or additionally be mediated via the child variables in Level 3.
- Level 3 includes variables which I consider to be personal characteristics or experiences of individual children, and which I hypothesise are 'closest' to children's mental health. Note that this Level includes several variables collected only in B-CAMHS99 or B-CAMHS04. Note also that the potential explanatory variables include internalising mental health problems but not prosocial behaviour. This is because of the evidence from Section 8.1 Chapter 8 that the prosocial SDQ subscale had poor discriminant validity relative to the externalising subscale.

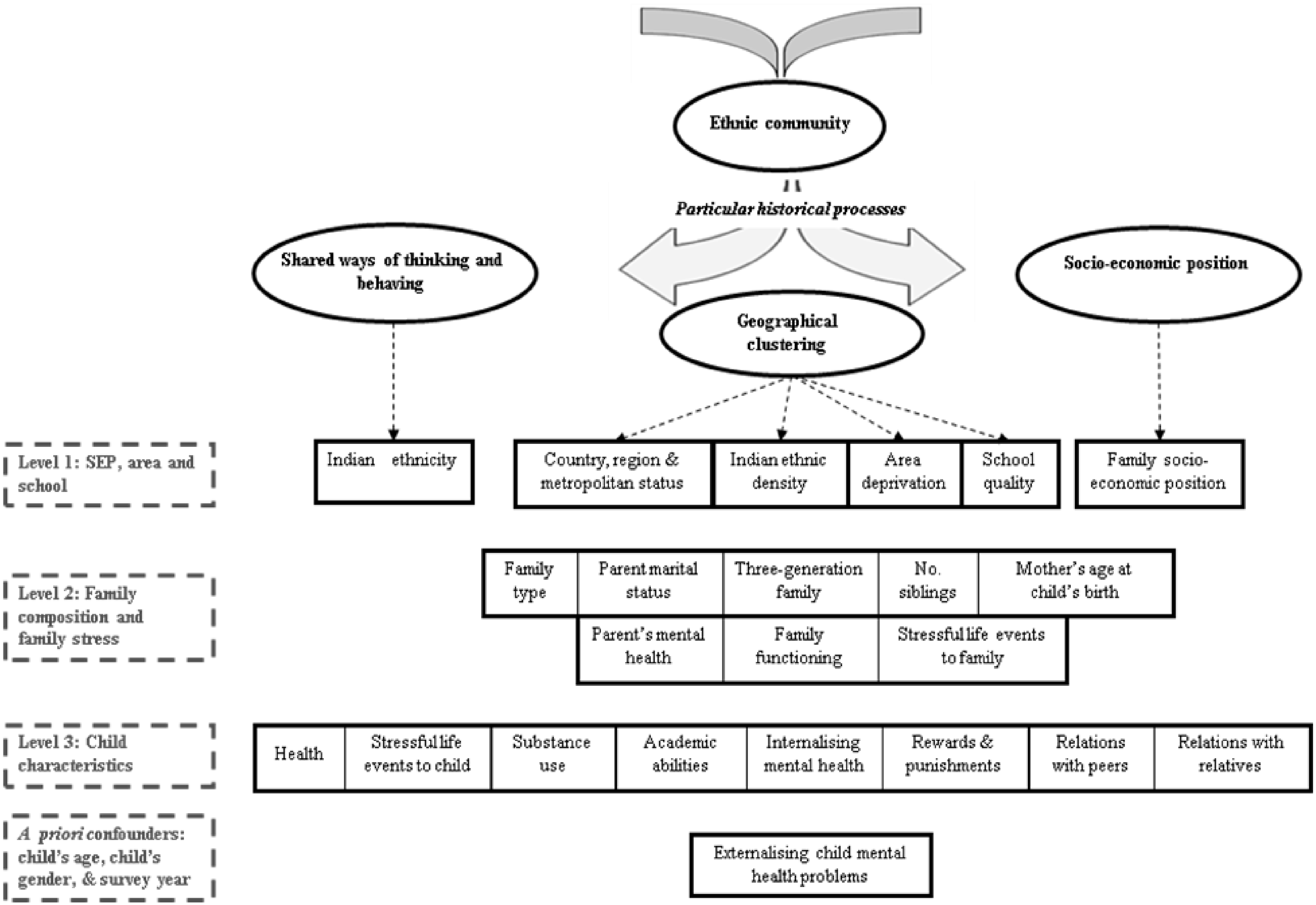
Finally, as in Chapter 7, I treat the child's age, gender and survey year (B-CAMHS99 vs. B-CAMHS04) as *a priori* confounders.

Table 9.1: Summary of explanatory variables in conceptual model

Exposure of interest	**Ethnicity** • Indian vs. White
A priori confounders	• Child's age • Child's gender • Survey year (B-CAMHS99 vs. B-CAMHS04)
Level 1: area characteristic, school characteristic and family socio-economic position	*Area characteristics* • Geographical area (country and region) • Metropolitan vs. non-metropolitan region • Indian ethnic density (proportion of local residents of Indian ethnicity) • Area deprivation (Indices of Multiple Deprivation) *School characteristics* • Ford Score (a predictor of the prevalence of mental health problems in a school) *Family socio-economic position* • Parental education • Household income • Housing tenure • Occupational social class • Mother's economic activity • Father's economic activity

Level 2: Family composition and family stress	*Family composition* • Family type (two-parent, stepfamily or lone parent family) • Parent marital status • Three-generation family • Number of co-resident siblings. • Mother's age at child's birth *Family stress* • Parent's mental health • Family functioning • Stressful life events affecting the whole family – Parental separation – Family financial crisis – Family member in contact with the police – Death of a parent or sibling
Level 3: Child characteristics	*Child characteristics* • Health – General health – Neuro-developmental disorder – Developmental problems or immaturity – Common physical health disorder – Rare physical health disorder • Stressful life events specific to the child – Illness requiring hospitalisation – Death of a friend • Substance use – Regular smoking – Alcohol consumption – Drug use • Academic difficulties – Teacher-reported academic difficulties – Parent-reported learning difficulties – Parent-reported dyslexia • Parent-reported internalising problems • Rewards and punishments (B-CAMHS99 only) – Rewards (praise; treats; favourite things) – Punishments (send to room; ground; shout; smack; ever hit/shake) • Relations with peers (B-CAMHS04 only) – Parent disapproves of friends – Parent thinks friends are trouble – Social aptitudes score • Relationships with relatives (B-CAMHS04, 11-16 year olds only) – Child's perceived social support – Number of relatives to whom child feels close – How often child helps relatives
Outcome	**Externalising problems** • Parent externalising SDQ score • Teacher externalising SDQ score

Figure 9.1: Hierarchical conceptual model for explanatory variables and externalising problems



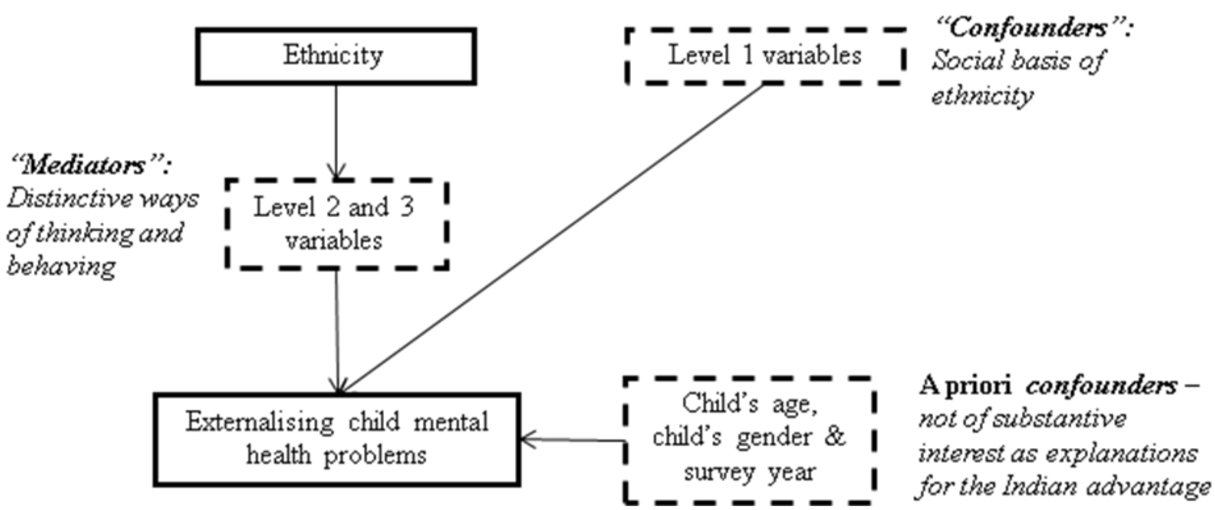
The relationship between ethnicity and the other explanatory variables

Section 3.1.1 Chapter 3 outlined the following theoretical framework for understanding ethnicity:

- **Biological bases for ethnicity**, based on allelic or epigenetic differences between groups.
- **Ethnic identity:** Ethnicity as a meaningful identity, created through sharing distinctive modes of speaking, thinking and behaving.
- **The social and structural basis of ethnicity**: Ethnicity as a social phenomenon generated and perpetuated by a group's structural position, including geographical clustering, socio-economic position and the experience of racism.

As discussed in Section 3.1.1, this thesis is guided by the hypothesis that the Indian mental health advantage results from aspects of the ethnic identity of British Indians and/or their social and structural position in British society. This is why Figure 9.1 places ethnicity in the structural/cultural Level 1 and not in the child-specific Level 3. Figure 9.2 illustrates how I conceptualise ethnicity as relating to the other explanatory variables, bringing together the above ethnicity typology and the concepts of mediating and confounding variables in causal models.

Figure 9.2: Conceptual model of the relationship between ethnicity and other explanatory variables: confounders, mediators and *a priori* confounders



Thus distinctive ways of thinking and behaving may generate differences between Indians and Whites in the Level 2 and 3 variables, and these may mediate the Indian mental health advantage. In addition, the Indian advantage may also be partly explained through other Level 1 variables, which represent part of the social basis of Indian ethnicity in Britain. I refer to these as ‘confounders’, which in standard epidemiological language would usually imply that they represent alternative, ‘non-causal’ explanations. In this instance, however, I believe a more nuanced interpretation is needed. This is because, as depicted at the top of Figure 9.1, historical processes related to ethnicity were central to the generation of the current geographical clustering and SEP of British Indians. These include the nature and timing of Indian migration to Britain and their subsequent patterns of structural assimilation (see Section 3.2 Chapter 3). As such, even if (for example) adjusting for geographical area of residence removed the Indian advantage entirely, it would not be correct to conclude that Indian ethnicity was unimportant. Rather I believe a better interpretation would be that the importance of ethnicity lay in the structural and social aspects of migration in the past, and not in the ethnic identity of Indians today.

A further note should be added about Indian ethnic density. This is the only variable in my conceptual model which was not envisaged as being a potential determinant of child mental health by the original B-CAMHS team. Indeed, I know of no reason to hypothesise that Indian ethnic density should be a predictor of child mental health in general. My inclusion of this variable stems instead from a desire to examine possible ethnic density effects. As described in Chapter 2 (p.39), these refer to the importance of the local concentration of an individual’s *own* ethnic group. The expectation is therefore that any effects of Indian ethnic density should differ between the White and Indian children in the sample, with effects in the latter being expected to be more pronounced. Indian ethnic density is consequently the only variable in my conceptual model which is not included because of its hypothesised main effect upon child mental health, but rather because of a hypothesised interaction with Indian ethnicity. I examine whether there is evidence of such an interaction in Section 11.1.1 Chapter 11.

To conclude, I consider all the Levels 1, 2 and 3 variables to be potentially important in explaining the Indian mental health advantage even though they differ in the hypothesised mechanisms of their effects. By contrast, I do not believe this applies to the three *a priori* confounders of child's age, gender and survey year. As argued in Chapter 7 (p.167), Indian ethnicity is not associated with these variables in the general population. It would therefore not be of substantive interest if the Indian advantage were partly 'explained' by an age, gender or survey year imbalance between Indians and Whites in B-CAMHS. In fact, as Chapter 7 reports, no such imbalances exist and confounding is therefore expected to be minimal.

9.2 Explanatory variables used in this PhD

I have previously described how B-CAMHS measured externalising problems (Section 5.3 Chapter 5) and ethnicity (Section 6.2.3 Chapter 6). I now describe the other child, family, school and area factors in my conceptual framework. Most variables were collected in both B-CAMHS surveys, and items were assessed through a mixture of verbal interviews with parents and children over 11 (with laptops for sensitive items); postal questionnaires to teachers; interviewer assessment; or using the child's postcode. This is summarised in Table 14.26 (Section 14.6.1 Appendix 2); for full copies of the B-CAMHS survey documents see [2-3].

9.2.1 Methods

Some of the variables described below had already been created by ONS, or else were very straightforward to generate from data already collected. Others required more complex calculation and evaluation, including comparisons of cross-cultural validity and/or psychometric performance between Indians and Whites.

Several child mental health risk factors in B-CAMHS were assessed using questionnaire measures. I compared these measures' factor structure between Indians and Whites using the same methodology described in Section 8.2 Chapter 8. Where *a priori* factor structures existed, I used multi-group confirmatory factor analyses (CFA) for ordinal data to make formal comparisons of model fit. Again, I first modelled first-order, second-order and general-specific CFA models in the pooled sample (i.e. Indians plus Whites) to assess which provided the best fit. Acceptable fit was defined as CFI>0.9, TLI>0.9, RMSEA<0.08 and all item factor loadings >0.4. I then used the best fitting model from the pooled-sample CFA in a multi-group CFA comparing Indians and Whites. Where no *a priori* factor structure was indicated by the literature, I identified factors using exploratory factor analyses (EFA) and a geomin rotation.

9.2.2 Level 1, Area characteristics

Geographical region and metropolitan areas

ONS used children's postcodes to assign them to Government Office Regions with Metropolitan counties. I grouped these regions in two ways (for details see Table 14.27 (Section 14.6.2, Appendix 2):

- **Geographical region**: An 11-fold geographical division, based upon England's eight Standard Statistical Regions (North East, North west, Yorkshire & Humberside, East Midlands, West Midlands, East Anglia, South East and South West), but with London additionally differentiated as a ninth category. Wales and Scotland were the tenth and eleventh categories.
- **Metropolitan area**: A binary division between metropolitan and non-metropolitan regions.

Indian ethnic density: proportion of residents in area of Indian ethnicity

I linked children's postcodes to 2001 UK census data on the ethnic composition of each child's Lower Super Output Area (LSOA). LSOA are geographical units of varying physical size but containing approximately 1500 individuals in the 2001 census [499]. I used the census data to calculate the proportion of inhabitants in each LSOA of Indian ethnicity. I could not calculate this variable for children in Scotland, because the Scottish census uses different geographic levels which do not include LSOAs.

Area deprivation: Indices of Multiple Deprivation (IMD)

I matched children's postcodes to their country's 2004/2005 Indices of Multiple Deprivation scores (2004 for England and Scotland, 2005 for Wales). The Indices of Multiple Deprivation (IMD) measure multiple domains of deprivation at the LSOA level [482]. Scores on these different domains can be used individually or combined into a weighted average to give an overall IMD score. The overall English IMD consists of 37 indicators in seven domains (income; employment; health and disability; education, skills and training; barriers to housing and services; crime; and living environment). The Scottish and Welsh IMD use somewhat different domains and indicators, although the theoretical underpinning and purpose is the same. Table 14.28 (Section 14.6.2 Appendix 2) provides full details of the English IMD's domains and indicator variables and justifies my choice of IMD over other measures of small-area deprivation.

9.2.3 Level 1, School characteristics

The Ford score

The Ford score predicts the prevalence of child mental health problems in mainstream schools using routine data collected annually by the English Office for Standards in Education (OFSTED). It is calculated as the weighted average of four variables: percentage pupils eligible for free school meals; percentage pupils with statemented special educational needs; school's unauthorised absence rate; school's exclusion rate. These indicators are not all routinely collected in Scotland and Wales.

The Ford score was developed by Ford *et al.* [500] using 7864 attending state-funded mainstream schools in England in B-CAMHS99. It was then validated by me using the 6445 English children in B-CAMHS04 whose schools had the relevant OFSTED data ([501], Appendix 3). Section 14.6.2 Appendix 2, gives details of the calculation and validation of the Ford Score.

9.2.4 Level 1, Family socio-economic position

In line with previous B-CAMHS analyses [2-3, 417], I use four main indicators of family socio-economic position (SEP): household income, the responding parent's education, housing tenure, and the occupational class of household reference person. I also use data on parental economic activity.

Household income

Parents were asked to give the household gross annual income from all sources (e.g. earned income, state benefit, pensions) and before deductions for national insurance or income tax. ONS then grouped their responses to give an ordered categorical variable and I coded this to reflect the mid-point of each category: £0-99 (coded 0.5); £100-199 (1.5); £200-299 (2.5); £300-399 (3.5); £400-499 (4.5); £500-599 (5.5); £600-769 (6.85); £770 or over (8.5).[26]

Responding parent's education

The parent participating in the B-CAMHS survey was asked for their highest educational qualification, and I used these as follows: 'No qualifications' (coded 1); 'Poor GCSEs, (grades D-F) or equivalent, including 'other' low-level qualifications' (2); 'Good GCSEs (grades A-C) or equivalent' (3); 'A-level or equivalent' (4); 'Diploma (e.g. in teaching or nursing) or equivalent' (5); and 'Degree level' (6).

Note that parent's education was collected only for the responding parent. It therefore usually refers to the mother but in around 5% of cases refers instead to the father or another parent informant. In Chapter 10 I return to the implications of this for my analyses.

Housing tenure

ONS used information on housing tenure to group households into home-owners (including houses with mortgages); households renting from the social sector; and households renting privately.

Occupational class of household reference person

Parents were asked to state their current/most recent job and (where applicable) that of their partner. ONS used this to generate the occupational social class of the household reference person ('head of household'). In B-CAMSH99 ONS used the Registrar General's Standard Occupational Classification (SOC) system of six social classes (I; II; III Non-manual; III Manual; IV; V) plus two additional categories of 'never worked' and 'full-time student'

[26] Throughout this Section, I give the coding for ordered categorical variables for which I later present mean scores.

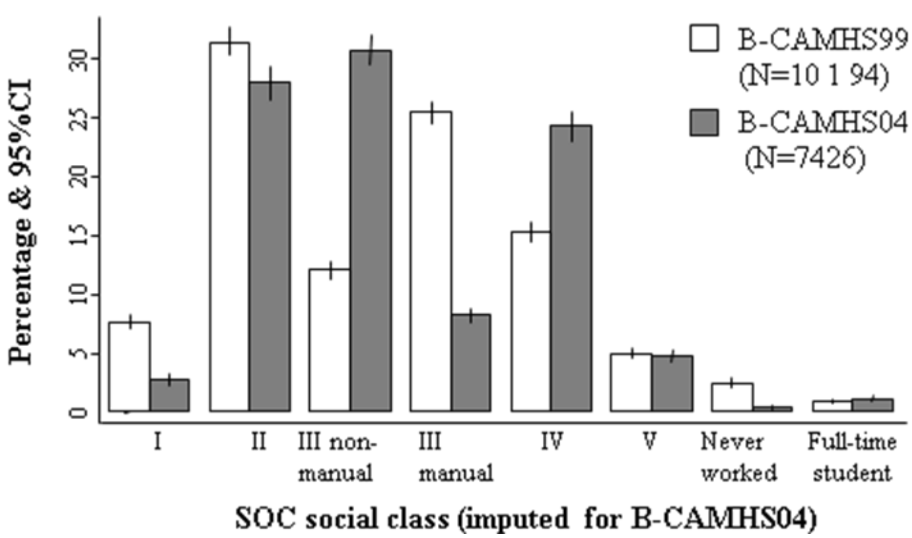
[502]. In B-CAMHS04 ONS used the 39 operational categories of the newly-created National Statistics Socio-economic Classification (NS-SEC) system [503].

Creating a combined measure of social class, and reasons for caution in its use

Of these two social class systems, the NS-SEC system has several conceptual and technical advantages, including a clearer theoretical underpinning and a more detailed classification scheme [504]. Unfortunately, the greater detail of NS-SEC means that one can translate from NS-SEC to SOC but not vice versa. I therefore converted the B-CAMHS04 NS-SEC codes into their approximate SOC equivalents. I used the translation algorithm which NS-SEC's creators developed using data from the UK census and other large surveys ([504, Appendix 2]; Table 14.30 Appendix 2).

The authors state that this translation algorithm achieves 87% continuity between SOC and NS-SEC [504]. Applying the algorithm did, however, generate a very different social class distribution in B-CAMHS04 compared to B-CAMHS99 (chi-squared $p<0.001$). In particular, there were many fewer individuals in social classes I and IIIM, and many more in social class IIIN (see Figure 9.3). This provides indirect evidence that the algorithm may be problematic. In Chapter 10 , I therefore conduct additional sensitivity analyses on the B-CAMHS04 sample in which I evaluate whether my substantive conclusions regarding the importance of occupational social class change when I use the ten standard NS-SEC analytic classes (Table 14.30 Appendix 2).

Figure 9.3: Relative frequency of different SOC categories by survey year



Mother's economic activity and father's economic activity

Parents were asked several questions about their current employment status and, where applicable, that of their partner. In conjunction with information about who was the responding parent (mother, father etc), I used this to create variables for 'mother's economic activity' and 'father's economic activity'. In lone parent families only one of these variables exists and in step-families one variable corresponds to a step-parent (e.g. 'father' in fact means 'stepfather').

The categories of economic activity which I used were 'full time employed', 'part-time employed', 'looking after home and family', 'unemployed' or 'other'. The category 'other' contained 4.2% of mothers and 5.4% of fathers, and included students, retired individuals, and individuals who were permanently or temporarily sick or disabled.

9.2.5 Level 2, Family composition

Following the standard census definition, ONS defined households as "a single person or group of people who have the accommodation as their only or main residence and who either share one meal a day or share the living accommodation" [3, p.386]. Parents were asked for the age, sex and marital status of all household members, and their relation to the index child. ONS used this to calculate:

- **Family type**, defined as 'two-parent' (containing the biological or adoptive mother and father); 'stepfamily' (at least one step-relative in the family); and 'lone parent family'.
- **Parent marital status**, defined as married vs. cohabiting. This variable did not exist for lone parent families.
- **Number of co-resident siblings** (including step-siblings) in the household.
- Mother's age at child's birth

I use all these variables, and also created an additional variable 'three generation family', which corresponded to any household in which a grandparent was living.

9.2.6 Level 2, Family stress

Parent's mental health

The 12-item version of the General Health Questionnaire (GHQ-12) was administered by laptop to parents. The GHQ-12 asks about depressive and anxiety symptoms over the past four weeks giving a score from 0 (no problem) to 12 (severe problem) [505]. It is probably the most widely used screening instrument for common mental disorders in community settings [506], and has been validated around the world [505, 507-508]. Particularly relevant to this thesis is its validation both in India [509-510] and in Indian-origin groups in Britain [511-512]. For the individual GHQ-12 items and fuller review of this existing literature, see Section 14.6.2 Appendix 2 (p.448).

Previous investigations have been somewhat inconsistent regarding the GHQ-12's factor structure [513-517], and I therefore applied an exploratory factor analysis (for ordinal data) to the B-CAMHS data. In both Indians and Whites there were two factors with an Eigenvalue of greater than one and these were very similar between the two groups. I used the factor structure indicated by the pooled sample as the basis for a multigroup CFA analysis. This showed adequate fit (CFI=0.983, TLI=0.987, RMSEA=0.070), indicating measurement invariance across Indians and Whites. Section 14.6.2 Appendix 2 provides full details of these analyses.

As with parent's education, parent's mental health was only measured for the responding parent – i.e. usually, but not always, the mother. Chapter 10 examines the implications of this for my analyses.

Family Functioning

The General Functioning (GF) subscale of the McMaster Family Activity Device was administered by laptop to parents. The GF scale is a 12-item measure of family functioning which generates an approximately continuous score between 1 (good family functioning) and 4 (poor family functioning) [518].

The GF scale has been shown to have good reliability, good predictive validity and an ability to distinguish high risk groups in various countries, including Britain [519-521]. It has also been used to investigate child mental health with apparent success in several cultural settings outside Western Europe [520, 522-524]. Nevertheless, there has been little rigorous cross-cultural evaluation of the GF scale, and I know of no relevant research in minority ethnic groups in Britain.

Most previous research has focussed on investigating the factor structure of the full Family Activity Device and not just the GF scale. An exploratory principle factor analysis indicated a two-factor structure in both Indians and Whites in B-CAMHS. These seemed to be tapping into valences rather than substantive constructs, with positively worded items forming one factor and the negatively worded items the other. In a multigroup CFA, a general-specific model of this factor structure showed evidence of measurement invariance between Indians and Whites (CFI=0.991, TLI=0.993, RMSEA=0.048). Section 14.6.2 Appendix 2 (p.449) gives details of these analyses, together with the constituent GF items and a more detailed review of evidence on its cross-cultural validity.

Stressful life events affecting the whole family

Parents were asked about a number of stressful life events, four of which would be expected to affect the whole family:

- Parental separation due to marital difficulties, or breaking off a steady relationship.
- Major financial crisis in the family.
- Family member [other than the child] having a problem with the police involving a court appearance.
- Death of parent, brother or sister of the child.

9.2.7 Level 3, Child characteristics

Health

General health

The parent was asked "How is (Child's) health in general?", choosing from the options 'very good' (coded 4), 'good' (3), 'fair' (2), 'bad' (1) or 'very bad' (0).

I know of no work which explicitly examines the cross-cultural validity of this item across ethnic groups. Yet s discussed in Chapter 3 (p.88), the 1999 Health Survey for England and other surveys have reported discrepant findings across ethnic groups between this general health measure and other outcomes such as limiting longstanding illness. This suggests the possibility of a reporting bias across ethnic groups, an issue which I explore empirically in Chapter 10.

Specific health complaints

The parent was asked to identify whether their child had any of a list of specific health complaints, described in detail in Section 14.6.2 Appendix 2 (p.451). I used these to create the following four binary variables:

- **Any specific neuro-developmental disorder**: epilepsy; cerebral palsy
- **Any non-specific marker of developmental immaturity or developmental disorder** ('developmental problems'): bed-wetting; speech and language problems; problems with coordination; muscle disease or weaknesses.
- **Any common physical disorder or complaint** (prevalence 2-15%): asthma; eczema; problems with eyesight; migraine; problems with hearing; glue ear, otitis media or grommits; food allergy.
- **Any rare physical disorder or complaint (prevalence<2%)**: stiffness or deformity of the foot, leg, fingers, arms or back; a heart problem; kidney, urinary tract problems; obesity; a condition present since birth such as club foot or cleft palate; diabetes; any blood disorder; cancer; missing fingers, hands, arms, toes, feet or legs; cystic fibrosis; chronic fatigue syndrome; spina bifida.

Note that as well as being rarer than the common physical disorders, the rare physical disorders will also usually be more serious.

Stressful life events specific to the child

In addition to the family-level stressful life events listed previously, parents were asked about a further two life events specific to the child:

- Child having a serious illness requiring a stay in hospital.
- Death of a close friend of the child.

There was also a third child-specific stressful life event, namely whether the child was 'in a serious accident or badly hurt in an accident'. I decided *a priori* to exclude this because the evidence reviewed in Chapter 2 (p.35) suggests that associations between accidents and externalising problems are most likely to reflect reverse causality. Only for traumatic brain injury is a 'forward' causal effect well-established and such injuries are rare. Indeed, head injuries of any sort seem to have been comparatively rare: questions on unintentional injuries in B-CAMHS99 found fractures to be by far the most common (lifetime prevalence of 15.8%), followed by head injury (4.3%), burns (2.3%), and poisoning (2.1%) [475]. Moreover, while some of these forms of injury were associated with ADHD there was no independent effect from head injuries [475].

Substance use

Children aged 11-16 were asked by laptop about their smoking, drinking and drug use. I used these to create three substance use variables:

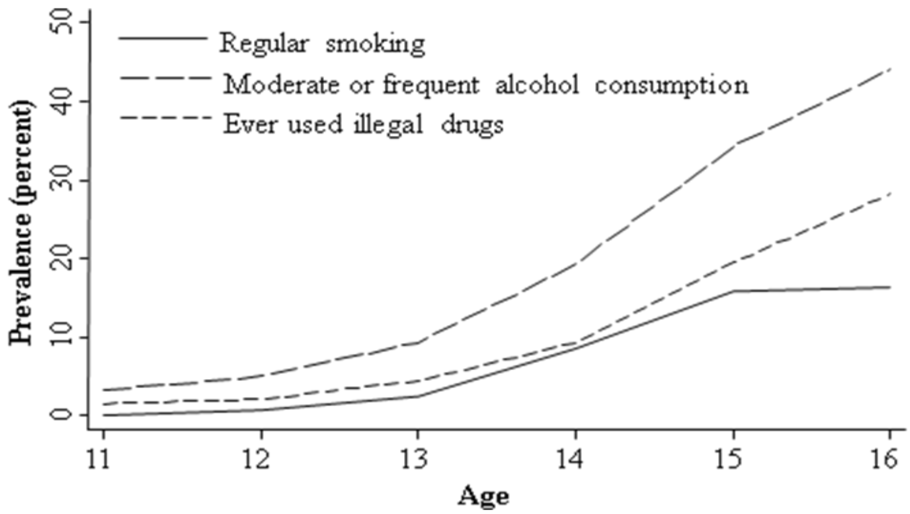
- **Regular smoker**, defined as smoking one or more cigarettes per week, in accordance with the standard ONS definition [525].
- **Alcohol consumption**, divided into never or rarely drinking alcohol (defined as less often than once a fortnight, coded 1), moderate alcohol consumption (drinking alcohol between once a week and once a fortnight, coded 2), and frequent alcohol consumption (drinking alcohol twice a week or more, coded 3).
- **Ever used illegal drugs**.

The smoking and illegal drug use variables are defined in the same way as in the RELACHS study in East London [346]. For alcohol consumption, RELACHS used a binary variable of 'drinking at least once a fortnight' but the authors note that a standardised definition does not exist for this age group. I therefore decided to use a more detailed categorisation of alcohol consumption at the high end of the range.

Information on smoking, drinking and drug use were only collected from children aged 11-16. I assumed all these behaviours were absent in children aged 5-10 years. I felt justified in this because at age 11 all these behaviours were rare to very rare, applying to just 0.1%

for regular smoking, 3.2% for moderate alcohol consumption (or 0.6% for frequent drinking), and 1.5% for drug use (see Figure 9.4).

Figure 9.4: Prevalence of substance use at ages 11 to 16 (N=7591)



Academic difficulties

The following measures of academic abilities/difficulties were available from one or both surveys (full details in Section 14.6.2 Appendix 2, p.452):

- Formal tests of general cognitive ability using the British Picture Vocabulary Scale, second edition, (BPVS-II [526]); formal tests of reading and spelling ability, using the British Ability Scales, second edition (BAS-II [527]). These three tests were administered in B-CAMHS99 only, to children of all ages.
- Parental assessment of the child's reading, maths, spelling and school work compared to the average, in B-CAMHS04 only
- Teacher assessment of the child's reading, maths, spelling and mental age compared to the average, in both B-CAMHS surveys.

I was unable to identify research investigating the validity of any of these measures across ethnic groups in Britain. In my own preliminary assessment of this issue, the only measure which appeared obviously problematic was the British Picture Vocabulary Scale (BPVS-II). There was strong evidence that Indians got worse scores on this measure than Whites ($p<0.001$) while the 10 other measures either showed an Indian advantage (seven measures at $p<0.05$) or a non-significant trend towards an Indian advantage (three measures). Even

Indian children with teacher-reported mental ages ahead of their chronological ages received BPVS-II scores well below the B-CAMHS average. I therefore excluded the BPVS-II from subsequent analyses. By contrast, the other academic measures showed a relatively consistent picture in which Indians were advantaged for spelling and maths, but perhaps not for reading. For fuller details of these analyses, see Section 14.6.2 Appendix 2 (p.452).

Teacher-reported academic difficulties

In light of the above evaluation, I decided to use the teacher assessments as my primary measure of academic difficulties in my substantive analyses. These had the advantages of being available from both surveys, of showing no evidence of cross-cultural bias, and of arguably being *a priori* less likely than parental assessments to show such cross-cultural biases.

The teacher assessments of the child's ability in reading, maths and spelling had response options were 'Above average' (coded 0), 'Average' (1), 'Has some difficulty' (2), 'Marked difficulty' (3). I summed these to create a score ranging from 0-9.

Sensitivity analysis using formal assessments of reading and spelling

The teacher-reported difficulties score has the advantage of being available for both datasets, but will plausibly show greater measurement error than the formal tests of reading and spelling ability. These formal tests also have the advantage of providing truly continuous measures of ability. I therefore decided to use these variables for a sensitivity analyses restricted to B-CAMHS99.

Parent-reported learning difficulties and dyslexia

Finally, both B-CAMHS surveys also asked parents two yes/no questions asked in about whether their child had 1) learning difficulties and/or 2) dyslexia.

Internalising problems

I used the parent internalising SDQ subscale as a measure of internalising problems. As demonstrated in Section 8.2 Chapter 8, this has good psychometric in both Whites and Indians. I chose the parent rather than the teacher or child internalising subscale because of the much higher data completeness from parents. Furthermore, as discussed in Chapter 2 (p.58), teachers typically provide less valid information for internalising problems.

Parent's use of rewards and punishments (B-CAMHS99 only)

In B-CAMHS99, parents were asked about their use of the following rewards and punishments. The response options to each question were 'Never' (coded 0), 'Seldom' (1), 'Sometimes' (2), 'Frequently' (3).

Rewards: "How often do you reward good behaviour or doing something well by…?"

1. Giving encouragement or praise?
2. Giving treats, such as extra pocket money, staying up late or a special outing?
3. Giving favourite things?

Punishments: "All children are naughty at sometime. How often do you punish (Child) when s/he misbehaves or does something wrong by…?"

1. Sending him/her to his/her room?
2. 'Grounding' him/her?
3. Shouting or yelling at him/her?
4. Smacking him/her with your hand?
5. Hitting him/her with a strap or something else?
6. Shaking him/her?

Most parents used the first four sorts of punishments at least occasionally, with only 5-55% of parents saying they 'Never' used each method. By contrast, the vast majority (98%) of parents said they 'Never' used hitting or shaking their child. I therefore combined these two rare and serious punishment categories into a single binary variable of 'ever hits or shakes child'.

Relationships with peers (B-CAMHS04 only)

Parent disapproves of friends

In B-CAMHS04, parents were asked whether they approved of the child's friends. The response options to which were 'A lot' (reverse coded as 0), 'A little' (1) and 'No' (2).

Parent thinks friends are trouble

In B-CAMHS04, parents were asked how many of the child's friends they thought were the sort of young people who often got into trouble. The responses were 'None' (coded 0) 'A few' (1) and 'Many' (2) or 'All' (3).

Social aptitudes scale

In B-CAMHS04, parents were asked to rate their children on a newly-created Social Aptitudes Scale (SAS). This consists of 10 statements scored on a five-point Likert scale, giving a total score of 0-40 (see Table 14.35, Section 14.6.2 Appendix 2 for individual items). Previous analyses in B-CAMHS04 have demonstrated high internal consistency between the scale's items (Cronbach alpha 0.88), with all the items loading heavily onto a single factor [528]. The Social Aptitudes Scale also had discriminant validity, being better than the SDQ total difficulty score at detecting autistic disorder but poorer at detecting emotional, behavioural and hyperactivity disorders.

Further exploratory factor analyses by me indicated that both Whites and Indians had one large Eigenvalue of over one (5.3 in Whites, 5.7 in Indians), on which all items loaded at >0.60 (see Table 14.35, Section 14.6.2 Appendix 2). A CFA of the single factor likewise showed evidence of measurement invariance between Indians and Whites (CFI=0.966, TLI=0.966, RMSEA=0.053). I therefore used the dimensional social aptitudes score in my analyses.

Relationships with relatives (B-CAMHS04 11-16 year olds only)

Perceived emotional social support

In B-CAMHS04, children were presented with seven statements about the emotional social support they received, with responses 'Not true' (coded 0) 'Partly true' (1) or 'Certainly true' (2) (individual items in Table 14.36, Section 14.6.2 Appendix 2).

The questions were taken from the 1985 Health and Lifestyle Survey (HALS) of 9003 adults in Britain [529]. I know of no previous research applying the questions to children, and they have been relatively little evaluated even in adults. One study provides a theoretical rationale for distinguishing received social support from anticipated social support, and reports that exploratory factor analyses support this two-factor solution in

HALS [530]. In B-CAMHS, however, exploratory factor analyses indicated only one factor with an Eigenvalue of >1, on which all seven items loaded strongly (loadings 0.65-0.83). It may be that in children the distinction between anticipated and received social support is not apparent because children are always so reliant upon care from others.

In Indians the factor loadings were somewhat lower for two items (see Table 14.36, Section 14.6.2 Appendix 2). This seemed likely simply to reflect the instability of estimates at small sample (N=86), however, as a CFA of the single factor showed evidence of measurement invariance between Indians and Whites (CFI=0.961, TLI=0.959, RMSEA=0.056). I therefore summed the responses from seven items to give a single score from 0-14.

Number of relatives to whom child feels close

In B-CAMHS04, children were asked two questions about the number relatives 1) in the household and 2) outside the household to whom the child felt emotionally close. Response options to both questions were ‘None’, ‘One’ or ‘Two or more’.

Helping relatives

In B-CAMHS04, children were asked how often they helped out relatives with activities such as shopping, cleaning or babysitting. The response options were ‘Less than once a month’; ‘At least once a month’; ‘at least once a week’; and ‘Every day’.

Chapter 10 Aim three, part two: Understanding the Indian advantage – descriptive, univariable and preliminary multivariable analyses

This Chapter defines the study population for my substantive analyses and compares the child, family, school and area characteristics of Indians and Whites. It then presents the cross-sectional association of each child, family, school and area characteristic with the parent and teacher externalising scores, and conducts preliminary analyses of how far these characteristics statistically explain the Indian mental health advantage. These analyses begin with 'complete case' univariable analyses and then move onto multivariable analyses with multiple imputation for missing data. The preliminary analyses in this Chapter are then refined in Chapter 11 Section 11.2, following additional investigations into the possibilities of interactions and reverse causality.

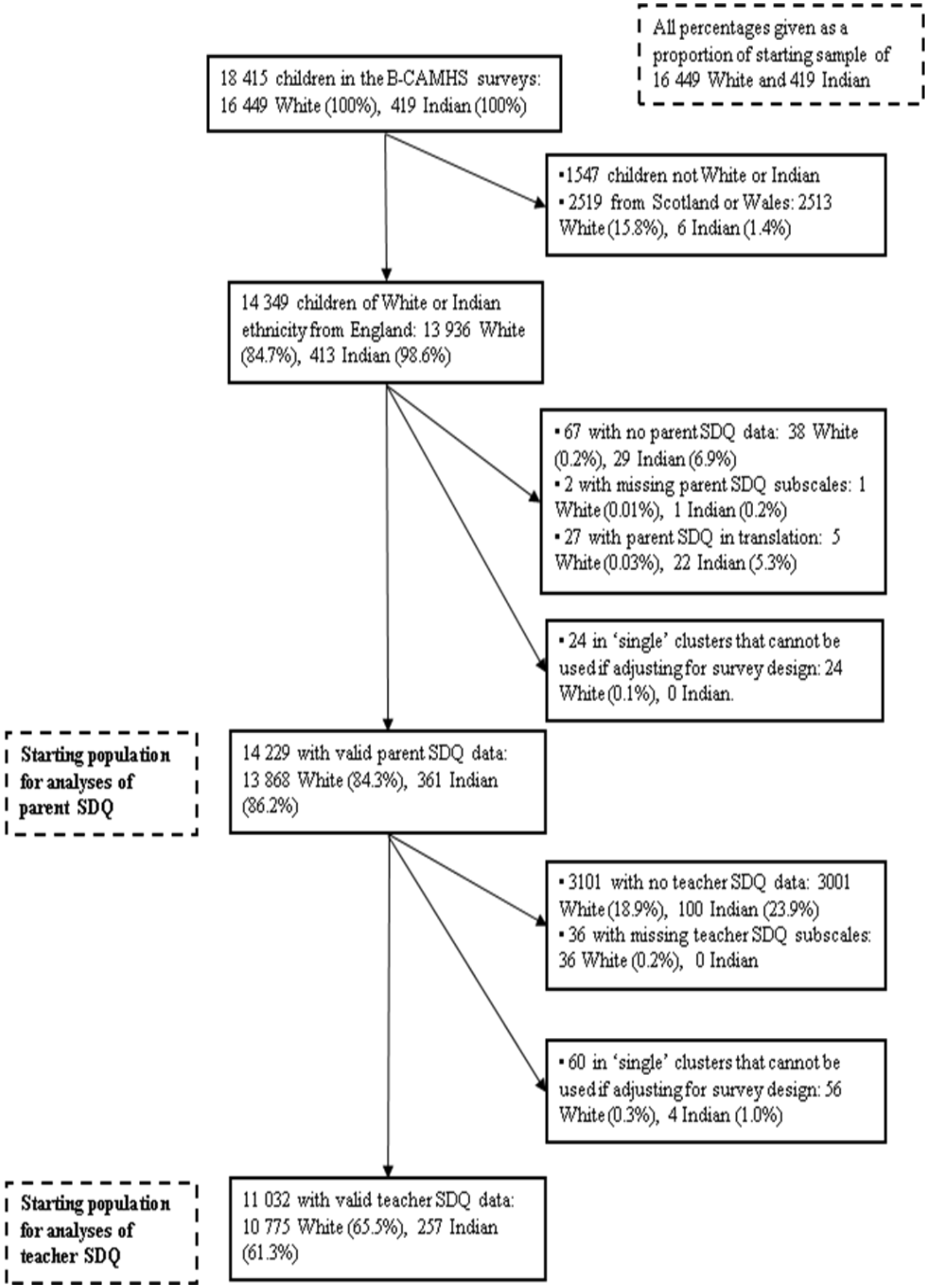
10.1.1 Methods

<u>Study population for substantive analyses</u>

I have decided to restrict my substantive analyses to children from England. This is partly because, as described in Section 9.2.2 and 9.2.3 Chapter 9, the Ford score, Indices of Multiple Deprivation and Indian ethnic density variables are not available in a comparable form in Scotland and Wales. Moreover, this restriction avoids any possibility of misleading findings resulting from some explanatory variables having different implications in England, Scotland and Wales, or there being different patterns of association between explanatory variables. As 413/419 of the B-CAMHS Indians lived in England, this restriction incurs very little loss of power.

The starting population for my substantive analyses is therefore the 14 229 English children (13 868 White, 361 Indian) with complete, non-translated parent SDQ data (Figure 10.1). Teacher SDQ data was available for 11 032 (77.5%) of these children (10 775 White, 257 Indian), and I seek to replicate all findings from the parent externalising score using the teacher-reported externalising score. Valid teacher SDQ data was also available for a further 64 children (23 White, 41 Indian) who did not have valid parent data, mostly because the parent had completed the SDQ in translation (11/23 Whites, 35/41 Indians). I decided not to include these 64 children because they were missing data for most family and child-variables (which were mostly obtained through parent-report) and therefore could not contribute much information to understanding the cause of the Indian mental health advantage. Moreover, I demonstrated in Chapter 8 (Table 8.19) that teacher-reported mental health did not differ between Indians whose parents answered the SDQ in translation and those whose parents answered in English. An advantage of excluding these children is to confine analyses of the teacher-reported externalising score to a stable subset of children from the parent analyses.

Figure 10.1: Starting population for analysis of parent and teacher externalising scores



Descriptive analyses of the characteristics of Indians and Whites

Mental health

I first compared graphically the distributions of the parent and teacher externalising scores for Indian and White children. I report the regression coefficient and significance of White (vs. Indian) ethnicity in linear regression analyses with parent/teacher-reported externalising score as the outcome and adjusting for the *a priori* confounders of age, gender and survey year.[27] I also tested for interactions between ethnicity and age or gender. In this and subsequent regression analyses, I verified that the assumptions underlying the regression models were met (see Chapter 6, p.149).

Child, family, school and area characteristics

I then compared the child, family, school and area characteristics of Whites and Indians, using the variables described in Section 9.2 Chapter 9 and summarised in Table 10.1. I report the proportion of missing data for each variable. Three variables only applied to some families (mother's economic activity, father's economic activity and parents married vs. cohabiting; see footnotes to Table 10.1). For these, I calculated missing data as a proportion of the relevant family types.

Among children who did have data, I calculated the percentage at each category/level of the explanatory variable and present a chi-squared test for association with ethnicity. For the purposes of these descriptive analyses, I categorised the continuous variables by creating bands of equal width with the aim of getting 4-7 bands in total (e.g. 2-point bands for the 12-point GHQ). If necessary, I used wider bands at the extremes to achieve at least 50 children (of either ethnicity) per category. I chose to use this categorisation approach rather than aiming for equal numbers per band (e.g. GHQ quintiles) because many variables had positively skewed distributions with few high-scorers. Dividing such distributions into equal numbers loses information about the tails [531].

For parent education, household income and all continuous variables I also present mean scores. I compared Indians and Whites using T-tests for variables that were approximately normally distributed, and Wilcoxon non-parametric tests for non-normal variables.

---

[27] This is the same method used to compare Indians and Whites in Chapter 7. Those analyses differed, however, in including children from Scotland and Wales or whose parents completed translated SDQs.

Table 10.1: Summary of nature of explanatory variables

Domain	Binary	Pure categorical	Ordered categorical	Continuous and scales
***A priori* confounders**	• Gender • Survey year			• Child's age
Area characteristics	• Metropolitan area	• Geographical region of England		• IMD score • Proportion Indian ethnicity
School characteristics				• Ford Score
Family SEP		• Housing tenure • Occupational social class • Mother's economic activity [nested†] • Father's economic activity [nested†]	• Parent's education • Household income	
Family composition	• Parents cohabiting vs. married [nested†] • Three-generational family	• Family type		• No. co-resident siblings • Mother's age at child's birth
Family stress	• Parent separation • Financial crisis • Family police contact • Death of parent or sibling			• Parent mental health • Family functioning
Child characteristics	• Neuro-developmental disorder • Developmental problems • Common physical disorder • Serious illness leading to hospitalisation • Death of friend • Rare physical disorder • Child smoking • Child drug use • Learning difficulties • Dyslexia • Punish: ever hit or shake (1999)		• Alcohol consumption • Reward: praise (1999) • Reward: treats (1999) • Reward: favourite things (1999) • Punish: send to room (1999) • Punish: grounding (1999) • Punish: shouting (1999) • Punish: smacking (1999) • Parent thinks friends trouble (2004) • Parent disapproves of friends (2004) • How often child helps relatives (2004)	• General health • Teacher-reported academic difficulties • Parent-reported internalising problems • Social aptitudes scale (2004) • Social support (2004) • No. close relatives in the home (2004) • No. close relatives outside the home (2004)

†Nested analyses: Mother's economic activity was only collected in households containing mothers (including adopted or stepmothers), and not in lone parent families headed by fathers or other family members. The equivalent was true for father's economic activity. Whether parents were cohabiting (vs. married) was not collected in lone parent families.
(1999)=collected in B-CAMHS99 only, (2004)=collected in B-CAMHS04 only

Inter-relations between explanatory variables

As discussed in Chapter 3 (p.83), the inter-relationship between SEP indicators may differ between minority ethnic groups and the White British population. I therefore explored this question in detail for the B-CAMHS family SEP and area deprivation variables. I also examined the association between other sets of indicators from the same domain (e.g. the physical health variables), in order to identify possible instances of reporting bias.

Sensitivity analyses of mother informants only

As described in Chapter 8 (Table 8.15), the proportion of responding parents who were fathers was higher in Indians than Whites. In the starting population for my substantive analyses, 13 098 (94.5%) of White parent respondents were mothers, 601 fathers (4.3%) and 169 'other' (1.2%). Among Indians, 297 were mothers (82.5%), 63 (17.5%) fathers, 0 'other', and one child had missing data. Parent informant type could therefore confound the relationship between ethnicity and parent education or parent mental health, these being the two variables which refer specifically to the responding parent. Moreover, while Section 8.3 Chapter 8 provided no evidence of systematic bias between mothers and fathers in reporting child mental health, such bias could exist for other parent-reported explanatory variables. To explore how far these concerns applied, I repeated the above descriptive analyses after restricting the sample to children with mother informants.

Association between externalising problems and child, family, school and area characteristics

I calculated the mean parent and teacher externalising scores for each category of the child, family, school or area factors, categorising continuous variables as described above. I did so firstly for the combined population of Indians and Whites, and then separately by ethnic group.

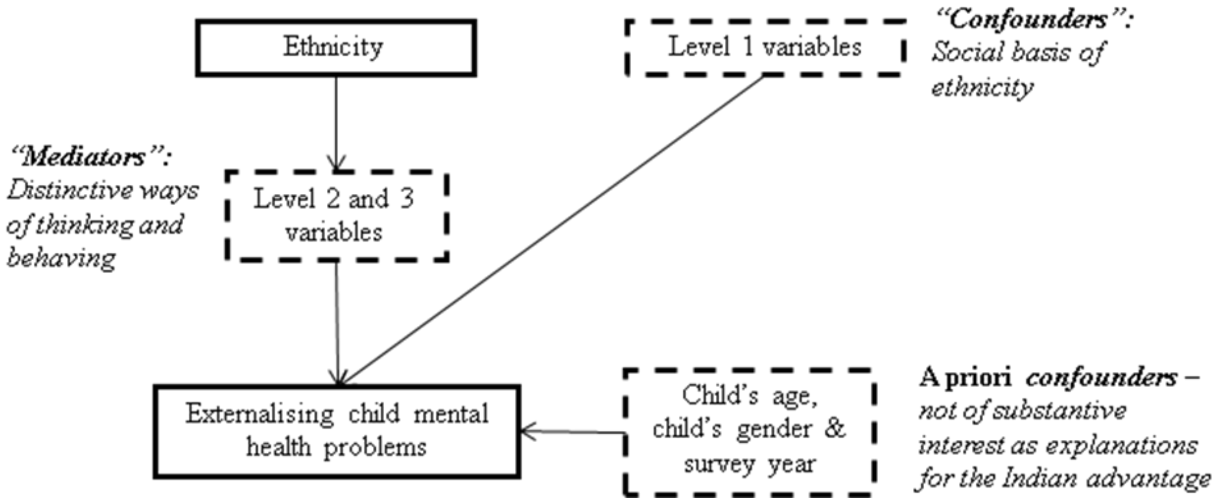
I present Wald p-values for the association in the combined population between each explanatory variable and the parent and teacher externalising scores. I calculated these p-values by fitting linear regression models with parent/teacher-reported externalising scores as the dependent variable and adjusting for age, gender and survey year. Into these models I entered in turn each explanatory variable listed in Table 10.1. I entered most explanatory variables as categorical variables except the continuous variables which I entered as linear terms. This had the advantages of increasing power for the continuous variables, reducing the potential for residual confounding through understratification, and avoiding decreasing precision by overstratification [532]. For continuous variables with more than five levels, I also assessed whether quadratic or cubic terms were significant predictors ($p<0.05$) of externalising problems. If so, then I indicate this and continued to include these quadratic/cubic terms in subsequent analyses. For continuous variables with five or fewer levels (general health, number of co-resident siblings, and number of close relatives) I did not assess quadratic and cubic terms. Instead I used likelihood ratio tests to compare models treating the variable as a linear term with models treating it as categorical, retaining the linear model if there was no evidence ($p<0.05$) that it provided inferior fit.

Finally, I tested for an interaction between parent informant type (mother, father or other) and parent education or parent mental health. These two variables were measured only in responding parents, and it seems plausible that their association with the child's externalising problems might vary across informant types. For example, maternal mental health might be more strongly associated with child externalising problems than paternal mental health. If so then adjusting for parental education without taking account of this interaction could give misleading findings when comparing Indians and Whites.

Effect of adjusting for child, family, school and area characteristics upon the magnitude of the Indian advantage: univariable analyses

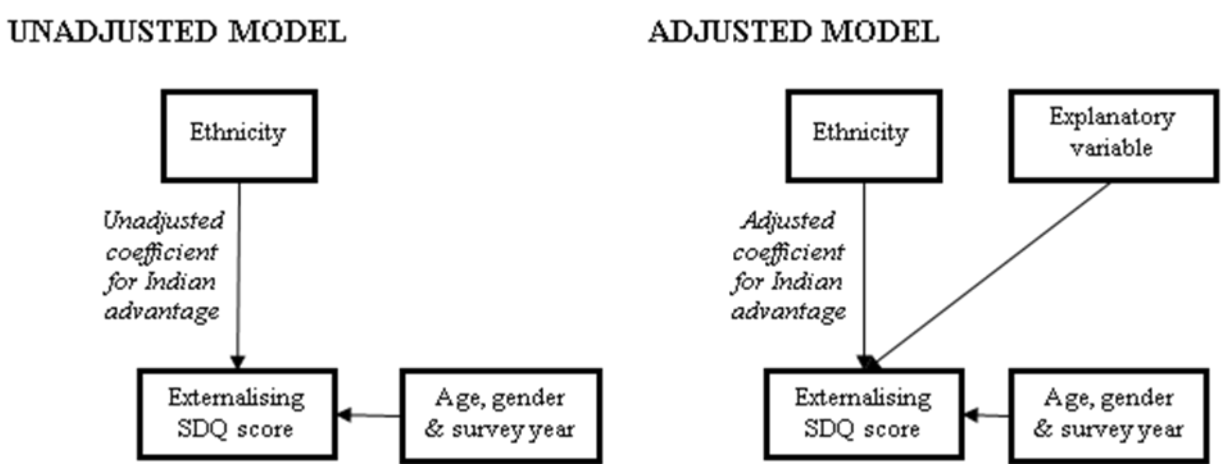
The main aim of this Chapter is to identify variables which are important in explaining the Indian mental health advantage, in accordance with the conceptual model presented in Chapter 9 and reproduced in Figure 10.2.

Figure 10.2: Conceptual model of the relationship between ethnicity and other explanatory variables: confounders, mediators and *a priori* confounders



As a preliminary assessment of this, I calculated how much the regression coefficient for White (vs. Indian) ethnicity changed after adjusting for each explanatory variable in turn (Figure 10.3). The expectation is that the more important a variable is as a confounder or mediator, the more adjusting for it should reduce the magnitude of the effect of ethnicity upon externalising problems.

Figure 10.3: Comparing the unadjusted and adjusted coefficients for the effect of ethnicity upon externalising problems



Note that explanatory variables from Levels 1, 2 and 3 were all entered as independent variables in the regression models, and that this statistical model cannot distinguish between confounders and mediators (see Figure 10.2).

I used linear regression models for these analyses, and always compared the adjusted and unadjusted coefficients between identical subsets of children. That is, I included in the 'unadjusted' models only children with data for the explanatory variable in question. I modelled explanatory variables in the same way (categorical, linear, linear plus quadratic etc) as when calculating univariable associations with externalising problems.

Multiple imputation for missing covariate data

Amount of missing data

The B-CAMHS surveys had little missing data, except as was missing systematically by B-CAMHS survey (i.e. only collected in one survey) or by teacher/child non-participation. For parent-reported variables collected in both surveys, the percentage of missing data was usually <1% and almost always <5%. The only exceptions were weekly household income (5.7% missing) and Ford score (9.9% missing). Teacher-reported variables were missing for the 22.5% of children with non-participating teachers and for a further 1-3% of children whose teachers did participate. Among 11 to 16 year olds, the child-reported substance use variables were missing for the 5% who did not participate and a further 1% of those who did.[28]

The reward and punishment variables were only collected in B-CAMHS99, and were therefore missing for 43% of the full sample. Parent-reported peer relations were only collected in B-CAMHS04 and were 57% missing. Child-reported relations with relatives were also only collected in B-CAMHS04 from 11 to 16 year olds and were 81% missing.

Multiple imputation model

To impute missing covariate values I used multiple imputation [533-534], using the MICE (multiple imputation by chained equations) command in Stata [535-536]. This imputes each missing value in the dataset in turn based on all other variables in the model, cycling through the variables and using both observed and imputed values with each subsequent cycle until all missing values are filled in. This represents one imputation. Multiple such imputations are then combined in order to model correctly the uncertainty inherent in

[28] As Chapter 9 (p.240) describes, I assumed regular smoking, moderate/frequent alcohol consumption and drug use were absent in children aged under 11.

imputing, rather than observing, data values. I used five imputations for these models. For more details regarding multiple imputation, see Section 13.4 Appendix 1.

This imputation method assumes data is missing at random, given the other variables in the imputation model. As described in full in Box 14.3, Section 14.7 Appendix 2, my imputation model included:

- All components of subsequent substantive models of interest, including outcome variables, explanatory variables and model structure (e.g. quadratic terms or potential interactions).
- All additional predictors of the variables with missing data.
- All predictors of data 'missingness' – i.e. predicting which individuals have missing data.

Effect of adjusting for child, family, school and area characteristics upon the magnitude of the Indian advantage: preliminary multivariable analyses

Correlation between explanatory variables

Given plausible levels of measurement error, it may be impossible to separate out the effects of strongly correlated explanatory variables [537]. To evaluate if this was a problem for these analyses, I calculated the correlation between each ordered categorical or continuous variable and all other ordered categorical/continuous variables. I used Pearson's coefficients for continuous-continuous and continuous-ordered categorical pairs, and Spearman's coefficients for ordered categorical-ordered categorical pairs.

Multivariable analyses

For multivariable analyses I used the dataset created by multiple imputation. For variables collected only in one dataset and/or only by child report, I restricted my analysis to that subpopulation (e.g. 'B-CAMHS99 only' or 'B-CAMHS04 11-16 year olds only'). This was to avoid using covariates for which over half of the values had been imputed, which would require strong assumptions regarding the validity of the imputation model and would generate large standard errors.

As in univariable analyses, my main interest was in how the coefficient for White (vs. Indian) ethnicity changed after adjusting for different sets of potential explanatory variables. First I adjusted for variables which in univariable analyses had reduced the regression coefficient for the parent-reported Indian advantage by 0.10 or more. I then additionally adjusted for variables which had changed the parent-reported coefficient by less than 0.10. Finally I additionally adjusted for variables which had increased the parent-reported coefficient by 0.10 or more, by which point all variables in the Level had been added. This strategy therefore included an 'extreme case' model calculating how close the coefficient reduced towards zero in the first model which adjusted only for those variables which in univariable analyses had moved the coefficient in that direction.

All multivariable models used forced entry to adjust for these variables, treating them as categorical or continuous according to how they were modelled in univariable analyses. In this Chapter, I conduct these analyses separately the three Levels in my conceptual model; I then present multi-Level models in Chapter 11.

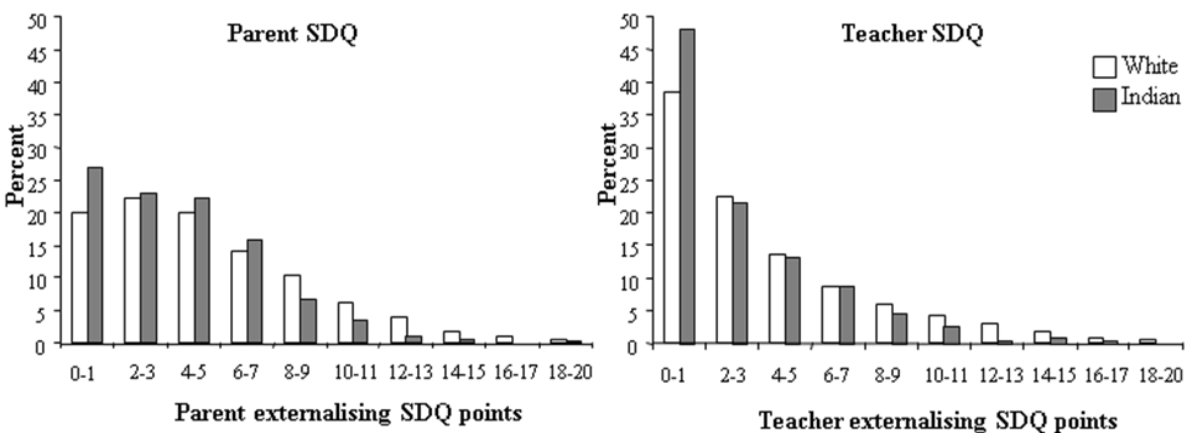
10.1.2 Results

Descriptive analyses of the characteristics of Indians and Whites

Externalising problems

There was strong evidence that Indian children had lower parent and teacher externalising scores (p<0.001). Moreover, the entire distribution of Indian externalising scores was shifted to the left, with more children receiving very low scores and fewer receiving very high scores (Figure 10.4 and Table 14.37, Section 14.7 Appendix 2). The Indian mental health advantage therefore applied across the whole range.

Figure 10.4: Parent and teacher externalising scores in Indians and Whites



Explanatory variables

As described below, many child, family, school and area characteristics showed major differences between Indians and Whites (full data presented in Table 14.37, Section 14.7 Appendix 2). These patterns were almost identical after restricting the analyses to children for whom the mother was the parent informant. The observed differences between Indians and Whites therefore cannot be attributed to confounding or bias resulting from the higher proportion of father informants among Indians.

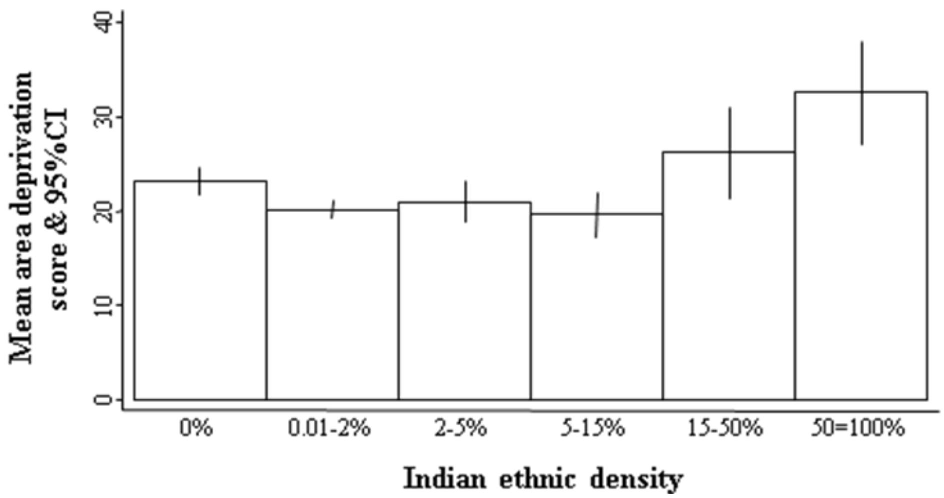
A priori confounders

There was no evidence (p>0.05) that ethnicity was associated with gender, age or survey year. The White sample was 50.8% male, had a mean age of 10.2 years and 42.0% were sampled in B-CAMHS04. Their Indian counterparts were 52.5% male, had a mean age of 10.3 years and 45.5% were sampled in B-CAMHS04.

Level 1 variables: area, school and family SEP

Indian children were concentrated in London and the Midlands, in metropolitan regions (69.3% vs. 44.4%), and in areas with a higher Indian ethnic density (mean ethnic density 23.2% vs. 1.3%). Indians also lived in more deprived areas (mean IMD score 26.9 vs. 21.0 in Whites, or 0.37 standard deviations). Area deprivation had a U-shaped relationship with ethnic density, being greater in areas containing 0% Indians or, particularly, in areas containing 15% Indians or more (Figure 10.5). This reinforces the importance of controlling for area deprivation when investigating the effects of ethnic density, as previously discussed in Chapter 2 (p.39). There was no evidence of a difference between Indians and Whites in the Ford score of their schools.

Figure 10.5: Mean area deprivation by Indian ethnic density



With regards to family SEP, income and social class distributions were relatively similar between Indians and Whites. For parental education, however, Indians were more concentrated at the extremes of the distribution, with more parents having both degree-level qualifications (18.6% vs. 12.5% in White) and no qualifications (28.3% vs. 19.8%).[29] Indians were also substantially more likely to be homeowners (88.7% vs. 71.0%) and less likely to be renting in the social sector (7.7% vs. 22.5%). These findings are consistent with the findings of other recent surveys, as reported in Section 3.2.2 Chapter 3.

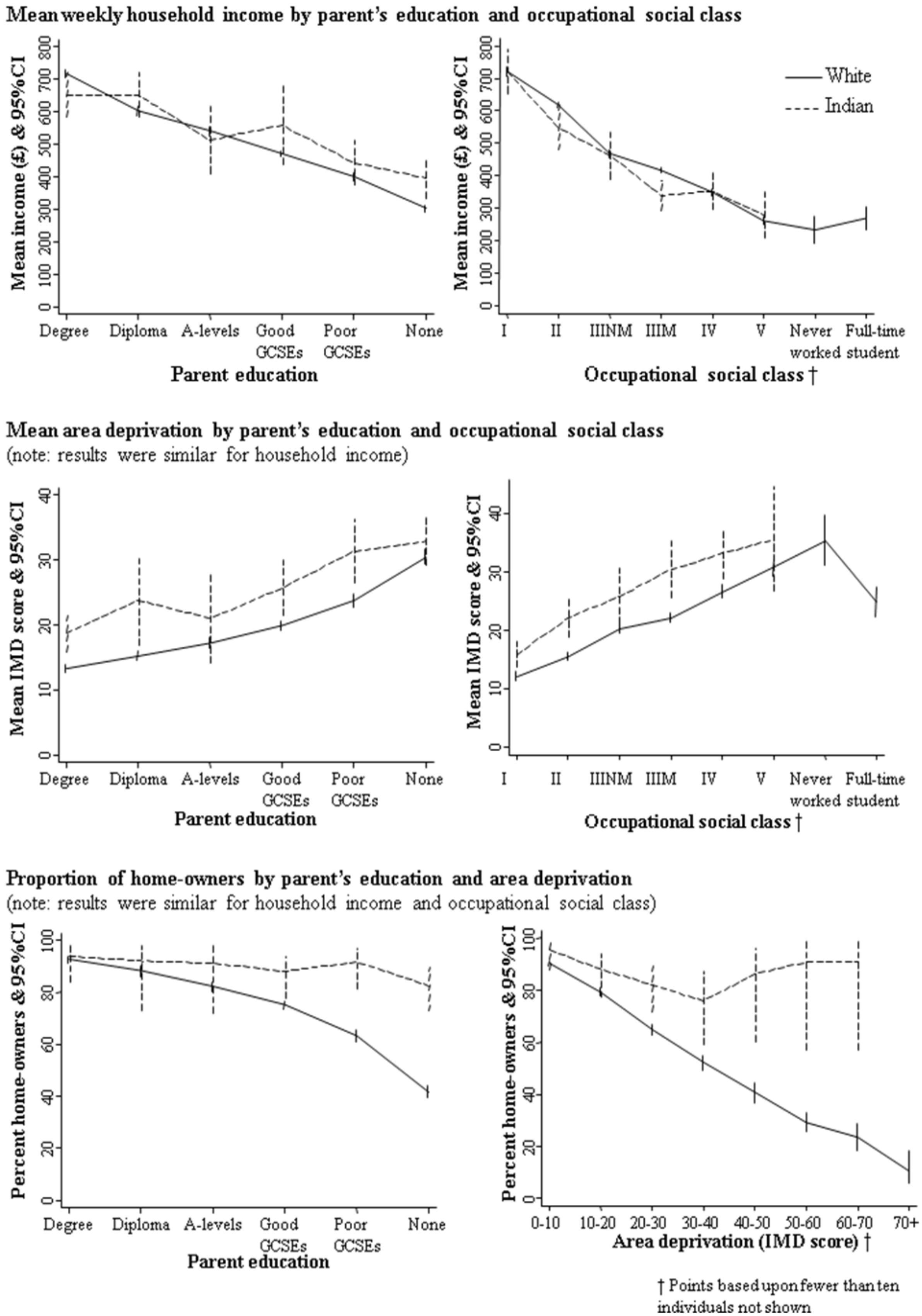
[29] Additional exploratory analyses indicated that this different distribution of parent education was not explained by the higher proportion of Indian father respondents. It did, however, seem partly to reflect many Indian parents immigrating to Britain when they were too old to take standard British secondary school exams. For details see Section 14.7, Appendix 2 (p.464).

Indians therefore appeared to be systematically disadvantaged for area deprivation, systematically advantaged in housing tenure, concentrated at the extremes of the distribution for parent education, and not much different for occupational social class and income. This mixed picture confirmed the need for a detailed examination of the pattern of inter-relationship between the different SEP indicators. Household income, parent education and social class showed a very similar relationship to each other in Indians and Whites. For example, as illustrated in Figure 10.6, household income was very similar in Indians and Whites after stratifying by parent education or social class. By contrast, area deprivation scores were systematically higher in Indians than Whites after stratifying by family SEP, but the gradient (i.e. the degree of social differentiation *within* ethnic groups) was similar. Home-ownership showed a different pattern again: the proportion of Indian and White home-owners was very similar in the most advantaged groups but whereas in Whites there was a steep gradient with SEP and area deprivation, this was not observed in Indians.[30]

Finally, Indian mothers were more likely to be working full-time (33.3% vs. 24.2% for Whites) or be looking after the home and family (31.2% vs. 23.4%), and less likely to be working part-time (28.3% vs. 46.0%). Indian fathers were somewhat less likely to be working full-time (79.6% vs. 87.3%), and instead slightly more likely to be working part-time, looking after home and family, unemployed or other.

[30] Note that 'difficulty of access to owner-occupation' (the modelled proportion of households unable in 2002 to afford to enter owner-occupation) forms part of the IMD. It is only one of 37 such indicators, however, and the circularity in comparing home-ownership with area deprivation is therefore low.

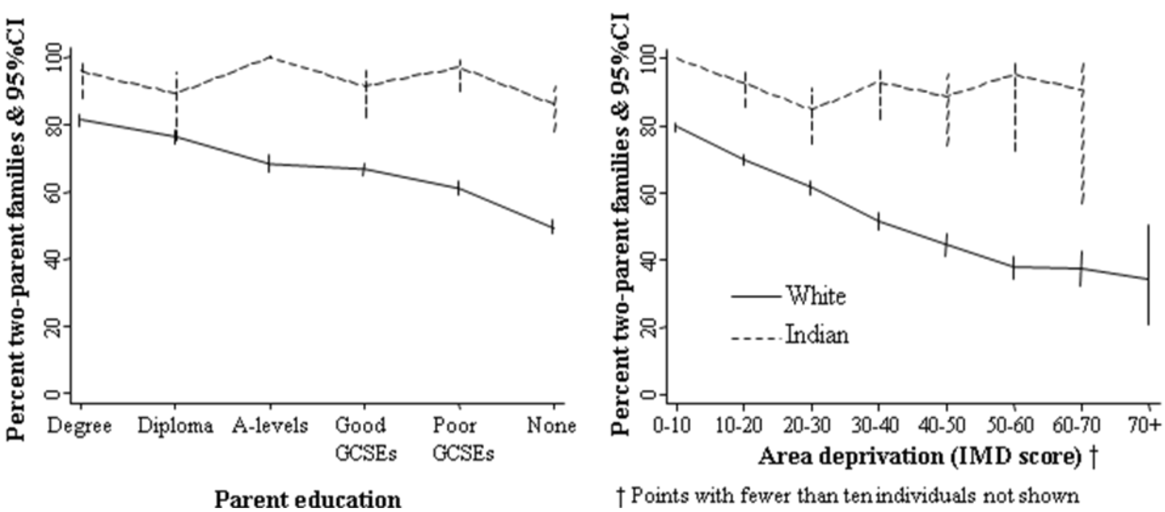
Figure 10.6: Interrelation between measures of family SEP and area disadvantage



Level 2 variables: family composition and family stress

Two-parent families were substantially more common in Indians (92.2% vs. 65.4% in Whites) and also less socio-economically differentiated (Figure 10.7). Indian parents were also more likely to be married rather than cohabiting (99.5% vs. 88.0%) and less likely ever to have experienced a parental separation (8.4% vs. 31.4%). Indian families were also more likely to have a grandparent in the household (14.5% vs. 1.9%) and contained slightly more co-resident siblings (mean 1.42 siblings vs. 1.27). There was no difference in the age of mother's at the time of the child's birth (mean 27.8 years vs. 27.9 in Whites).

Figure 10.7: Prevalence of two-parent families in Indians and Whites, by parent education and area deprivation



Note: the pattern was similar when other SEP indicators were used.

Among the family stress variables, there was no evidence of an ethnic difference in parent mental health (mean score 1.75 vs. 1.71)[31] or in the likelihood of the death of a family member. Indian families were, however, less likely to have experienced a financial crisis (10.6% vs. 15.1%) or to have had a family member in trouble with the police (3.1% vs. 6.2%). On the other hand, there was also strong evidence that Indian families had worse parent-reported family functioning (mean 1.80 vs. 1.69, or 0.27 standard deviations).

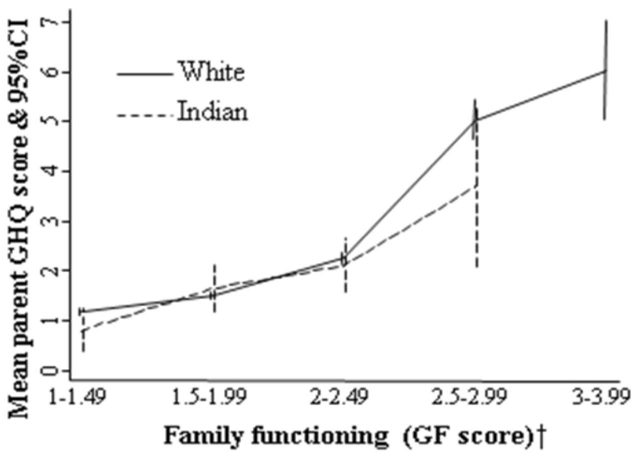
This pattern of ethnic similarities and differences in family composition and family stress was almost unchanged after restricting the analyses to two-parent families. The only

[31]This might seem inconsistent with the mental health advantage in Indian children. It should be remembered, however, that the GHQ measures parental depression and anxiety – i.e. precisely the emotional mental health problems for which Indian children likewise showed no difference from Whites.

substantive changes in analyses of two-parent families were that Indians no longer had lower police contact (2.8% vs. 2.9% in Whites) but Indian mothers were now younger (mean age 27.8 years at the child's birth, vs. 28.9 years in Whites). This latter difference resulted from younger White mothers being particularly under-represented among two-parent families: only 47% of White mothers who had their child aged under 25 were living in a two-parent family at the time of the B-CAMHS interview, as compared to 75% of White mothers aged 25 or more. By contrast, the proportion of Indian mothers in two-parent families was 93% for both age categories.

Family type therefore did not seem to explain the different profiles of Indians and Whites for the other Level 2 family variables. Because the worse parent-reported family functioning in Indian families was unexpected, I investigated further its relationship with parent mental health. I hypothesised that these two family stress indicators would be positively associated. This was indeed the case. Moreover, the strength of the correlation was very similar in the two ethnic groups (Pearson's coefficient 0.25 in Indians and 0.26 in Whites), and the mean GHQ score of Indian and White parents was similar after stratifying by family functioning (Figure 10.8). This provides some evidence that the GF scale provides a measure of family functioning which is comparable between Indians and Whites. This is in accordance with the demonstration in Chapter 9 (p.236) that it had a very similar factor structure in the two groups.

Figure 10.8: Mean parent GHQ score in Indians and Whites, stratified by family functioning



† Points with fewer than ten individuals not shown

Level 3 variables: child characteristics

The parent-reported physical health data was inconsistent in terms of whether it indicated better or worse health in Indians than Whites. Indian parents provided less favourable assessments of their children's general health than White parents (54.1% Indian children in 'very good health' vs. 70.3% Whites). Yet Indian parents also reported their children had fewer developmental problems (4.7% vs. 9.7% in Whites), fewer serious illnesses requiring hospitalisation (11.3% vs. 17.8%), fewer common physical complaints (33.6% vs. 39.6%) and fewer rare physical complaints (3.3% vs. 6.4%).

This discrepancy prompted me to investigate further the relationship between these different measures of physical health. One would expect parent-reported general health to decline with increasing numbers of physical complaints, and this was observed indeed in both ethnic groups (Figure 10.9). At any given number of physical complaints, however, Indian parents rated their children's general health less favourably than White parents. In linear regression analyses predicting general health and adjusting for number of physical complaints, age, gender and survey year, the mean general health score of Whites was 0.24 points higher than Indians (95%CI 0.17, 0.32), corresponding to 0.38 standard deviations. This value was almost unchanged after adjusting for each specific disorder separately (difference 0.23, 95%CI 0.16, 0.31). A similar pattern emerged in further analyses of parent-reported general health after stratification by the 'somatic problems' item on the teacher and child SDQ (Table 10.2). Again, both Indians and Whites children had poorer parent-reported general health if the teacher or child had reported somatic symptoms, but within each stratum the Indian mean was about 0.2 points lower.

Within Indians and Whites, the relationship between the different measures of physical health is therefore as expected, indicating that these are not meaningless measures in either group. There is, however, a systematic difference between the groups in parent-reported general health. This is not explained by specific physical disorders and is also observed after stratifying by teacher- and child-reported somatic symptoms. This may therefore represent a reporting bias in the general health variable, whereby White parents systematically make more favourable assessments than Indian parents. If so, adjusting for general health may generate misleading findings when seeking to explain the Indian advantage. I return to this issue in the next Chapter.

Figure 10.9: Parent-reported general health by number of physical complaints in Indians and Whites

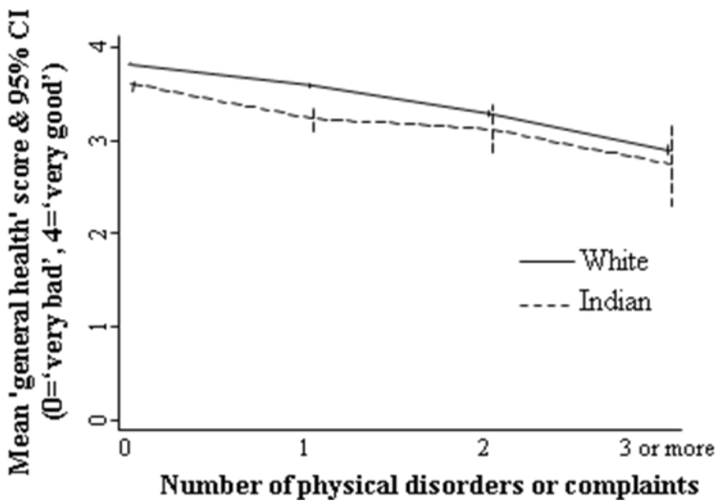


Table 10.2: Parent-reported general health for Whites and Indians, stratifying by teacher- and child-reported somatic symptoms

	Response to somatic symptoms SDQ item†	**White**		**Indian**		**P-value for ethnic difference**
		N	**Mean (95%CI)**	**N**	**Mean (95%CI)**	
Teacher-report	**'Not true'**	8982	3.69 (3.67, 3.70)	234	3.45 (3.37, 3.54)	<0.001
	'Partly/ certainly true'	1824	3.43 (3.40, 3.47)	28	3.32 (3.04, 3.60)	0.44
Child-report	**'Not true'**	3585	3.74 (3.72, 3.76)	101	3.52 (3.38, 3.66)	0.002
	'Partly/ certainly true'	2153	3.46 (3.43, 3.50)	53	3.26 (3.09, 3.43)	0.02

†"Often complains of headaches, stomach-aches or sickness"

Indian parents reported a lower prevalence of learning difficulties (2.9% vs. 8.6% in Whites) and dyslexia (0.5% vs. 3.6%) in their children, and this was supported by some evidence for fewer teacher-reported academic difficulties in Indians (mean 2.71 vs. 3.03, or 0.13 standard deviations). There was no evidence of cross-cultural bias for these academic variables; rather the mean teacher-reported academic scores were closely similar between Whites and Indians within each stratum (Table 10.3). Note, however, that the very small number of Indians with a learning difficulty or dyslexia (N=9) means that this analysis is highly underpowered in this stratum.

Table 10.3: Teacher-reported academic difficulties in Whites and Indians, stratifying by parent-reported learning difficulties and/or dyslexia.

	White		Indian		p-value for ethnic difference
	N	**Mean (95%CI)**	**N**	**Mean (95%CI)**	
Neither learning difficulty nor dyslexia present	9461	2.63 (2.58, 2.68)	243	2.54 (2.23, 2.85)	0.59
Learning difficulty and/or dyslexia present	1102	6.44 (6.32, 6.58)	9	6.83 (5.42, 8.24)	0.60

Indian young people had substantially lower alcohol consumption (1.6% drinking moderately or frequently, vs. 8.0% in Whites). There was no evidence of a difference in smoking or drug use, which were rare in both groups, although the trend was again for these to be less common in Indians. As in Chapter 7, there was no evidence of a difference between Indians and Whites for internalising problems.

The B-CAMHS99 data on parental rewards and punishments presented a rather mixed picture. Almost all parents (≥98%) in both ethnic groups praised their children 'frequently' or 'sometimes', but Indian parents were substantially less likely to praise their children 'frequently' (62% vs. 83.4% of Whites). Indian parents also gave treats slightly less frequently, but there was no difference for giving favourite things. Among the punishments, sending a child to their room or grounding them were less common in Indian parents, and 'frequent' shouting was also less common. The proportion of parents who 'never' smacked their child was slightly higher in Indians (61.6% vs. 55.0%), but so too was the proportion who smacked their child 'sometimes' or 'frequently' (14.8% vs. 9.2% in Whites). Moreover, while rare in both groups, ever hitting or shaking the child was more commonly reported by Indians (7.9% vs. 2.6% in Whites)

Finally, in B-CAMHS04 Indian children were no different in terms of parent disapproval of friends. They were, however, less likely to have friends whom their parents thought were 'trouble' (20.7% vs. 34.6%) and had better social aptitudes (mean 26.6 vs. 24.5, or 0.33 standard deviations). With regard to relationships with relatives, Indian children reported helping their relatives somewhat more often (e.g. 28.0% 'every day', vs. 14.8% in Whites), but there was no evidence of an ethnic difference in perceived emotional social support or in the number of relatives to whom the child felt close.

Association between externalising problems and child, family, school and area characteristics

Table 10.4 presents the mean parent externalising scores for each explanatory variable in the full Indian and White sample, and separately by ethnic group. As described in more detail below, most child, family, school and area factors showed strong evidence of an association with parent externalising scores (almost always p<0.001). This association was observed across the range in most continuous and ordered categorical variables. Even when quadratic or cubic terms were significant, the relationship between the explanatory variable and externalising problems was still usually monotonic rather than U-shaped or with any large threshold effect. Table 14.39 (Section 14.7 Appendix 2) presents the same analyses for the teacher-reported externalising scores. In almost all cases the substantive findings were identical, with any exceptions indicated in the text below.

Table 10.4: Cross-sectional association between parent externalising score and child, family, school and area characteristics

Domain	Variable	Categories	Mean parent externalising score			
			Full sample	p-value	White mean	Indian mean
***A priori* confounders**	**Child's sex**	Male	5.63	<0.001[a]	5.66	4.25
		Female	4.25		4.27	3.51
	Child's age	5-6 years	5.45	<0.001[b]	5.47	4.67
		7-8 years	5.23		5.26	4.00
		9-10 years	5.00		5.03	4.13
		11-12 years	4.75		4.78	3.85
		13-14 years	4.71		4.74	3.64
		15-16 years	4.39		4.43	2.87
	Survey year	1999	5.1	<0.001[a]	5.1	4.4
		2004	4.8		4.8	3.3
Area	**Geographical region**	South East	4.85	0.05[a]	4.86	3.67
		London	4.73		4.84	3.61
		South West	4.85		4.85	[3.92]
		Eastern	4.75		4.77	2.37
		East Midlands	4.97		4.99	4.56
		West Midlands	5.13		5.17	4.17
		North East	5.31		5.31	[3.98]
		North West & Merseyside	5.08		5.10	4.20
		Yorkshire & Humberside	5.06		5.08	2.62
	Metropolitan region	Non-Metropolitan	4.94	0.85[a]	4.94	4.28
		Metropolitan	4.97		5.02	3.73
	Indian ethnic density	<0.01%	5.08	<0.001[c]	5.08	[4.19]
		0.01-2%	4.97		4.98	3.33
		2-5%	4.85		4.91	3.66
		5-15%	4.46		4.55	3.62
		15-50%	4.39		4.59	4.19
		50-100%	4.61		[10.63]	4.21

Domain	Variable	Categories	Mean parent externalising score			
			Full sample	p-value	White mean	Indian mean
	Area deprivation	0-10 points	4.24	<0.001[c]	4.25	3.77
		10-20 points	4.60		4.62	3.61
		20-30 points	5.26		5.30	3.82
		30-40 points	5.77		5.87	4.02
		40-50 points	6.18		6.23	5.06
		50-60 points	6.30		6.40	3.61
		60-70 points	6.14		6.25	3.49
		70+ points	6.04		6.05	[5.00]
School	**Ford score**	0-2	4.00	<0.001[b]	4.01	3.66
		3-5	4.72		4.73	3.76
		6-8	5.27		5.30	4.43
		9-11	6.09		6.16	3.64
		12-14	6.35		6.42	[3.98]
		15-17	6.87		6.95	[3.00]
Family SEP	**Parent's highest educational qualification**	No qualifications	6.23	<0.001[a]	6.30	4.27
		Poor GCSEs	5.57		5.62	4.02
		Good GCSEs	4.84		4.86	4.08
		A-level	4.44		4.45	3.90
		Diploma	4.14		4.15	3.86
		Degree	3.47		3.49	3.06
	Weekly household income	£0-99	6.19	<0.001[a]	6.23	[3.97]
		£100-199	6.28		6.30	5.19
		£200-299	5.72		5.81	3.71
		£300-399	5.30		5.31	4.78
		£400-499	4.95		4.96	4.60
		£500-599	4.48		4.49	3.68
		£600-769	4.21		4.22	3.51
		£770 and over	3.79		3.80	3.64
	Housing tenure	Owner occupied	4.38	<0.001[a]	4.40	3.84
		Social sector rented	6.62		6.64	4.37
		Privately rented	5.58		5.60	4.21
	Occupational social class	I	3.77	<0.001[a]	3.77	3.91
		II	4.18		4.19	3.78
		III Non-manual	4.93		4.96	3.62
		III Manual	5.19		5.21	4.33
		IV	5.66		5.71	4.00
		V	6.20		6.22	4.94
		Never worked	7.15		7.16	[6.19]
		Full-time student	5.55		5.55	[empty cell]
	Mother's economic activity [nested]	Full-time employed	4.59	<0.001[a]	4.63	3.56
		Part-time employed	4.57		4.58	4.15
		Home and family	5.79		5.85	4.27
		Unemployed	5.62		5.68	[2.75]
		Other	5.65		5.75	2.99
	Father's economic activity [nested]	Full-time employed	4.53	<0.001[a]	4.54	3.91
		Part-time employed	4.88		4.92	4.18
		Home and family	5.94		6.06	[3.28]
		Unemployed	6.52		6.59	5.31
		Other	5.58		5.73	2.54
Family composition	**Family type**	Two-parent family	4.43	<0.001[a]	4.45	3.84
		Step family	6.04		6.05	[3.86]
		Lone parent family	5.92		5.93	4.70

Domain	Variable	Categories	Mean parent externalising score			
			Full sample	p-value	White mean	Indian mean
	Marital status [nested]	Married	4.52	<0.001[a]	4.55	3.84
		Cohabiting	5.85		5.85	[5.04]
	Three generation family	No grandparent in household	4.94	0.24[a]	4.97	3.82
		Grandparent in household	5.23		5.41	4.36
	Number of co-resident siblings	0	4.75	<0.001[a]	4.78	2.93
		1	4.80		4.82	3.98
		2	4.99		5.01	4.19
		3	5.87		5.93	4.30
		4 or more	6.22		6.32	3.61
	Mother's age at child's birth	≤19	6.45	<0.001[c]	6.49	4.04
		20-24	5.81		5.87	3.93
		25-29	4.79		4.81	4.16
		30-34	4.42		4.44	3.52
		35-39	4.05		4.07	3.46
		40 or more	4.41		4.41	[4.48]
Family stress	**Parent mental health**	0-1	4.51	<0.001[d]	4.53	3.78
		2-3	5.47		5.50	4.10
		4-5	5.60		5.66	3.31
		6-7	6.21		6.26	4.56
		8-9	6.39		6.43	[4.29]
		10-12	6.85		6.91	4.83
	Family functioning	1.0-1.49	3.92	<0.001[c]	3.93	3.08
		1.5-1.99	4.86		4.88	3.84
		2.0-2.49	5.90		5.95	4.47
		2.5-2.99	8.20		8.33	4.70
		3.0-4.0	8.68		8.68	[empty cell]
	Parental separation	No	4.49	<0.001[a]	4.51	3.87
		Yes	5.96		5.98	4.14
	Family financial crisis	No	4.86	<0.001[a]	4.89	3.87
		Yes	5.42		5.44	4.12
	Family police contact	No	4.83	<0.001[a]	4.85	3.92
		Yes	6.67		6.72	3.29
	Death of parent or sibling	No	4.92	<0.001[a]	4.94	3.89
		Yes	5.66		5.69	[4.08]
Child	**General health**	Bad/very bad	7.82	<0.001[b]	7.96	[4.88]
		Fair	7.06		7.15	4.43
		Good	5.79		5.85	4.52
		Very good	4.47		4.49	3.38
	Neuro-developmental disorder	No	4.92	<0.001[a]	4.95	3.89
		Yes	8.02		8.03	[6.00]
	Developmental problems	No	4.68	<0.001[a]	4.70	3.83
		Yes	7.56		7.59	5.27
	Common physical disorder	No	4.63	<0.001[a]	4.66	3.61
		Yes	5.44		5.46	4.47
	Rare physical disorder	No	4.88	<0.001[a]	4.91	3.88
		Yes	5.95		5.97	4.57

Domain	Variable	Categories	Mean parent externalising score			
			Full sample	p-value	White mean	Indian mean
	Serious illness leading to hospitalisation	No	4.80	<0.001[a]	4.82	3.90
		Yes	5.63		5.66	3.81
	Death of friend	No	4.89	<0.001[a]	4.92	3.87
		Yes	5.75		5.76	4.66
	Regular smoker	No	4.84	<0.001[a]	4.86	3.95
		Yes	7.49		7.51	[5.75]
	Alcohol consumption	Never/rare	4.90	<0.001[a]	4.93	3.94
		Moderate	4.87		4.87	[4.80]
		Frequent	5.43		5.42	[8.00]
	Ever used drugs	No	4.86	<0.001[a]	4.89	3.93
		Yes	6.07		6.08	[5.38]
	Teacher-reported difficulties in school	0-1	3.17	<0.001[b]	3.17	2.96
		2-3	4.62		4.64	3.63
		4-5	5.79		5.81	5.32
		6-7	6.91		6.93	5.88
		8-9	8.41		8.45	6.30
	Learning difficulty	No	4.59	<0.001[a]	4.61	3.81
		Yes	8.86		8.88	6.77
	Dyslexia	No	4.87	<0.001[a]	4.90	3.86
		Yes	7.06		7.04	[12.12]
	Parent-reported internalising SDQ score	0-1	3.50	<0.001[b]	3.52	2.58
		2-3	4.50		4.52	3.59
		4-5	5.46		5.50	4.15
		6-7	6.68		6.70	5.92
		8-9	7.69		7.72	6.32
		10-11	8.00		8.10	5.68
		12-13	9.97		10.14	[6.26]
		14-15	10.35		10.44	[2.00]
		16-17	11.71		11.86	[6.00]
		18-20	11.04		11.04	[empty cell]
Child, 1999 only	**Reward: praise**	Never	7.16	<0.001[a]	7.26	[5.00]
		Seldom	6.86		6.87	[6.65]
		Sometimes	5.99		6.02	5.45
		Frequently	4.89		4.91	3.67
	Reward: treats	Never	4.28	<0.001[a]	4.35	2.97
		Seldom	4.64		4.65	3.96
		Sometimes	4.96		4.98	4.54
		Frequently	5.47		5.49	4.54
	Reward: favourite things	Never	4.71	0.05[a]	4.73	[3.63]
		Seldom	4.87		4.88	4.62
		Sometimes	5.04		5.05	4.63
		Frequently	5.32		5.35	3.79
	Punish: send to room	Never	4.11	<0.001[a]	4.14	3.40
		Seldom	4.44		4.43	5.00
		Sometimes	5.86		5.88	5.14
		Frequently	7.93		7.95	[5.97]
	Punish: grounding	Never	4.09	<0.001[a]	4.09	4.12
		Seldom	4.77		4.76	5.70
		Sometimes	6.37		6.42	3.98
		Frequently	8.83		8.83	[8.26]

Domain	Variable	Categories	Mean parent externalising score			
			Full sample	p-value	White mean	Indian mean
	Punish: shouting	Never	3.81	<0.001[a]	3.86	2.66
		Seldom	3.92		3.93	3.49
		Sometimes	4.77		4.77	4.61
		Frequently	6.69		6.72	4.81
	Punish: smacking	Never	4.40	<0.001[a]	4.42	3.98
		Seldom	5.63		5.65	4.39
		Sometimes	6.94		6.97	6.15
		Frequently	10.12		10.57	[4.63]
	Punish: ever hit or shake	Never	5.03	<0.001[a]	5.05	4.25
		Ever	7.28		7.39	5.83
Child, 2004 only	**Parent disapproval of friends**	Approves a lot	4.25	<0.001[a]	4.28	3.05
		Approves a little	6.88		6.98	4.36
		Does not approve	9.50		9.52	[8.00]
	Parent thinks friends are trouble	None are trouble	3.92	<0.001[a]	3.94	3.22
		A few are trouble	5.90		5.94	3.53
		Many are trouble	10.44		10.49	[6.00]
		All are trouble	10.99		10.99	[empty cell]
	Social aptitudes score	0-9	13.15	<0.001[c]	13.15	[empty cell]
		10-14	11.06		11.10	[5.00]
		15-19	7.14		7.22	4.22
		20-24	4.93		4.96	3.88
		25-29	3.69		3.72	2.68
		30-34	3.07		3.05	3.45
		35-40	1.95		1.94	2.05
	Social support score	0-7	6.30	<0.001[c]	6.30	[empty cell]
		8-9	5.97		5.96	[6.36]
		10-11	5.86		5.94	3.80
		12-13	4.68		4.69	4.23
		14	3.41		3.43	2.29
	No. close relatives in the Home	None	6.27	<0.001[a]	6.36	[3.00]
		One	5.24		5.27	[3.90]
		Two or more	3.92		3.94	3.24
	No. close relatives outside the home	None	4.29	<0.001[a]	4.35	2.94
		One	4.93		4.97	[2.14]
		Two or more	4.02		4.03	3.65
	How often child helps relatives	Every day	4.23	0.12[a]	4.30	2.89
		At least once a week	4.08		4.08	3.95
		At least once a month	4.17		4.21	2.70
		Less than once a month	4.40		4.40	[4.21]
		Never	5.10		5.13	[0.00]

Nested analyses: Mother's economic activity was only collected in households in which the mother (or mother substitute) was present; father's economic activity where the father was present; and marital status where both were present. †[a]=variable entered as categorical; [b] variable entered as a linear term; [c] variable entered as linear plus quadratic terms; [d] variable entered as linear, quadratic and cubic terms. Cells in square brackets are based on fewer than 10 children.

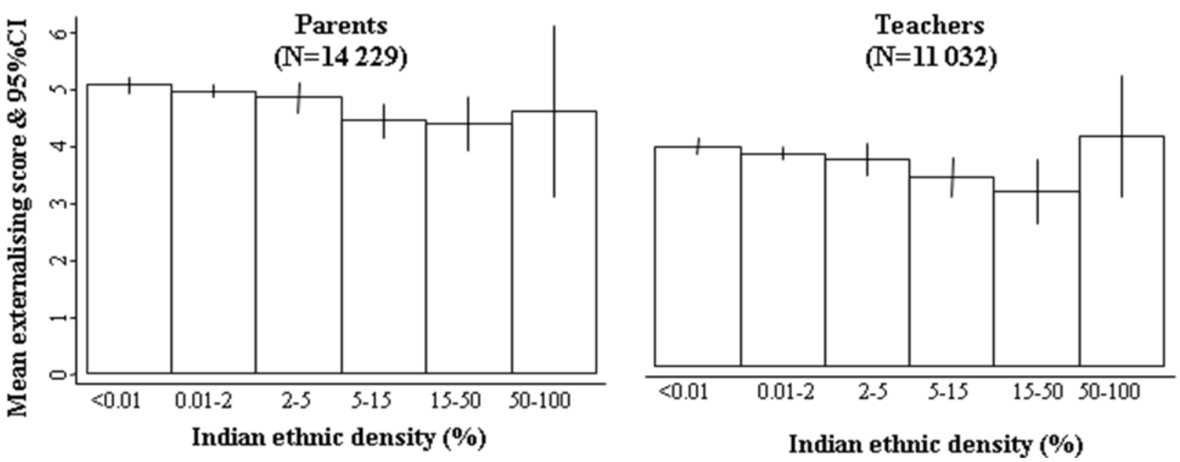
A priori confounders

Parent and teacher externalising scores were higher in boys and younger children. As in Chapter 7 (p.168), likelihood ratio tests provided no evidence ($p>0.60$) for either the parent or teacher-reported externalising scores that treating age as a categorical variable (using one-year age bands) provided a better model fit than treating it as a continuous variable. I therefore adjust for age as a continuous variable in all subsequent analyses. There was also no evidence ($p>0.05$) of an interaction between ethnicity and age or ethnicity and gender in the effects upon externalising problems.

Level 1 variables: area, school and family SEP

There was little or no evidence of any difference in externalising scores by geographical region or between metropolitan and. non-metropolitan regions. Externalising scores were, however, substantially higher in more deprived areas and schools with higher Ford scores. Externalising scores were also somewhat higher in areas of low Indian ethnic density and, by teacher report in areas of high Indian ethnic density (Figure 10.10). This was the only instance of any potential explanatory variable showing an apparent U-shaped relationship with externalising scores, a point to which I return in Chapter 11. In Chapter 11 I also examine whether, in line with the ethnic density hypothesis, the effects of Indian ethnic density differed between Indians and Whites.

Figure 10.10: Parent and teacher externalising scores by Indian ethnic density



With regard to family SEP, externalising scores were high in families in which parents were unemployed or at home looking after the family, as opposed to working full- or part-time. They were also higher in families with lower parental education, lower incomes, rented housing, and lower occupation social class. There was weak evidence ($p=0.03$) of an interaction between parent's education and parent informant type (mother vs. father vs. 'other'). This seemed likely to represent a chance finding, however, as it was not replicated when the outcome was the teacher-reported externalising score (p=0.38), child-reported externalising score (p=0.28) or externalising DAWBA diagnosis (p=0.78).

Level 2 variables: family composition and family stress

Externalising scores were higher in step- or lone parent families compared to two-parent families. Externalising scores were also high in families with cohabiting (vs. married) parents, with three or more co-resident siblings and with younger mothers. Parental mental health problems, poor family functioning and all the stressful life events were also associated with higher externalising scores. There was no evidence of an interaction between parent mental health and parent informant type (p=0.36 for the parent externalising score, p=0.10 for the teacher score).

The parent and teacher externalising scores showed conflicting associations with living in a three-generation family. There was evidence that children in three-generation families had higher teacher-reported externalising scores (p=0.008) but no evidence of an association with parent scores (p=0.24). There was likewise no evidence of an association with the child-reported externalising score (p=0.72) or externalising DAWBA diagnosis (p=0.66). The association between living in a three-generation family and the teacher-reported externalising score therefore seems likely to be a chance finding.

Level 3 variables: child characteristics

Mean externalising scores were higher among children with poorer physical health (by any measure), who had experienced any stressful life event, who used any substance, who had more academic difficulties (by any measure) and who had more internalising problems.

All five punishment variables showed a positive association between greater frequency of punishment and more externalising problems, with this being observed across the range. By contrast, there was an interesting and unanticipated heterogeneity between the reward strategies in their associations with externalising scores. Greater frequency of praising children was associated with fewer externalising problems but, contrary to what might be expected, the reverse was true of giving of treats and (at borderline significance) giving favourite things. A plausible explanation for the positive association between giving treats/favourite things and externalising problems is that parents use these reward strategies to deal with difficult children. I return to this issue of reverse causality for this and other child and family variables in the next Chapter.

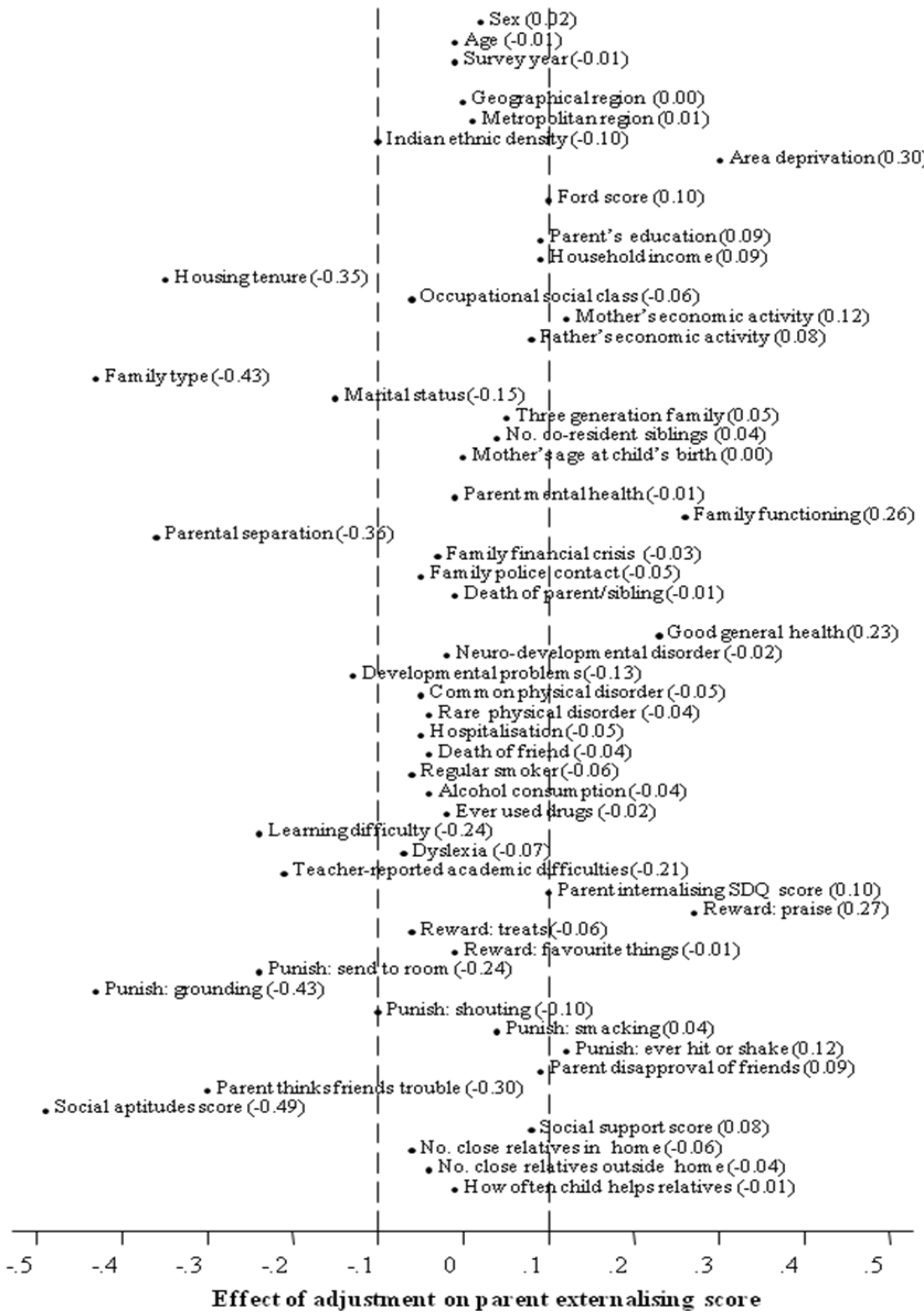
Externalising scores were higher among children who had more friends of whom their parent disapproved, who had more friends who were trouble or who had poorer social aptitudes. Scores were also higher in children who reported themselves to have lower social support or fewer close relatives living in the home. As for close relatives living outside the home, children with one such relative had higher externalising scores than children with either none or with two or more. This may be because having precisely one close relative outside the home is a marker for parental separation, the relative in question being the non-resident parent. Certainly 48.9% of children who reported having one close relative outside the home came from lone or step-parent families, as compared to 29.7% of those with no close relatives outside the home and 34.7% with two or more close relatives outside the home (p-value for heterogeneity <0.001). There was no evidence of an association between externalising problems and the frequency of helping one's relatives.

Effect of adjusting for child, family, school and area characteristics upon the magnitude of the Indian advantage: univariable analyses

The regression coefficient for the difference between Whites and Indians on the parent externalising score was 1.08 (95%CI 0.73, 1.43) after adjusting for age, gender and survey year. That is, the externalising score of Whites was an average of 1.08 points higher (less favourable) than Indians, corresponding to 0.28 standard deviations. The corresponding difference for teacher scores was 1.05 (95%CI 0.67, 1.43) or 0.26 standard deviations.

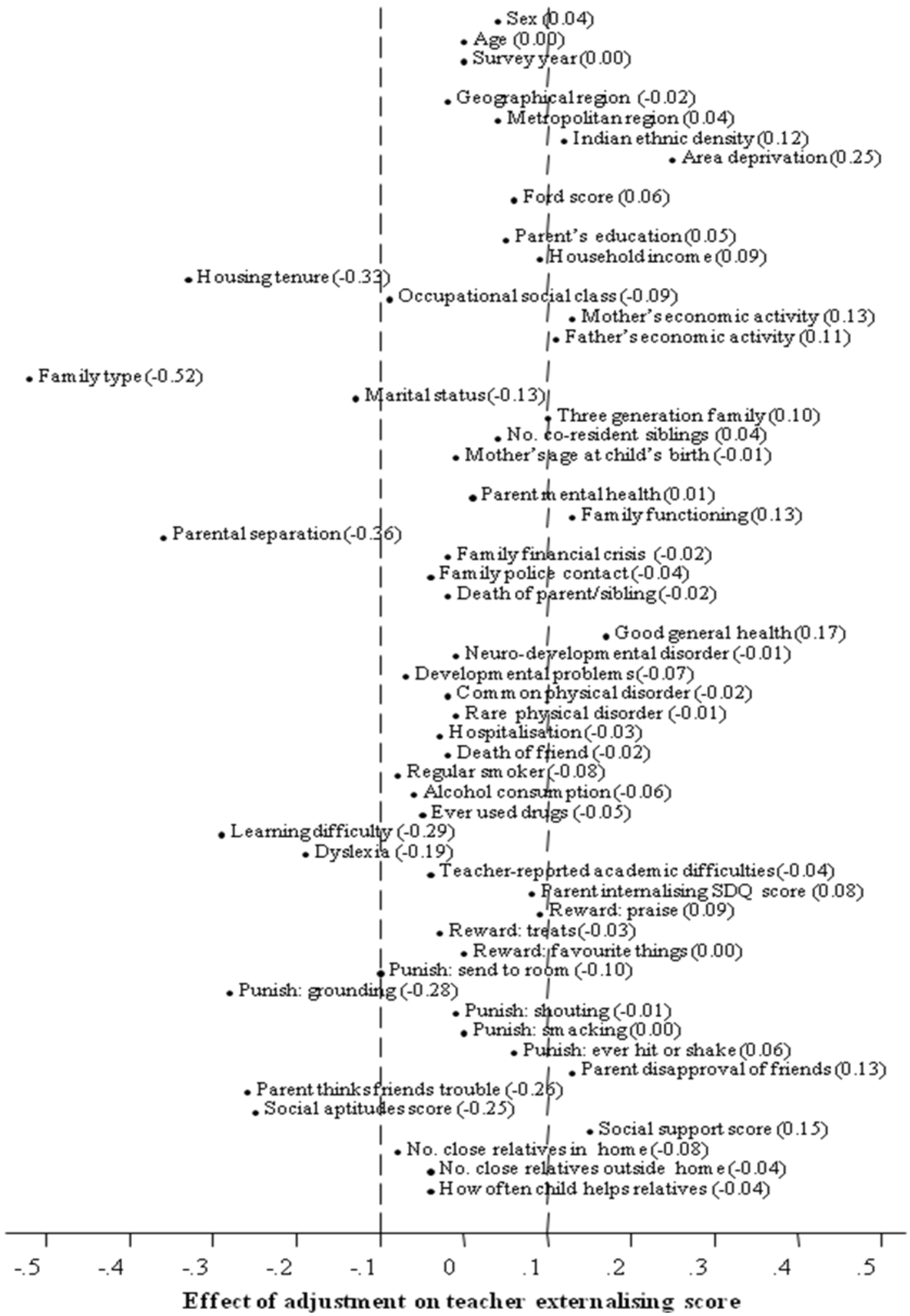
I recorded how this regression coefficient changed after additionally adjusting for each potential explanatory variable in turn. I interpreted reductions in the magnitude of the regression coefficient as corresponding to the Indian advantage being 'explained' and increases in the regression coefficient as corresponding to the Indian advantage being 'unmasked'. The results are summarised graphically in Figure 10.11 and Figure 10.12, and presented in full in Table 14.40 and Table 14.41, Appendix 2. For most variables the parent and teacher results were very similar, with any exceptions indicated below.

Figure 10.11: Effect of adjusting for child, family, school and area characteristics upon parent externalising score (complete case analysis)



For full results, see Appendix 2, Section 14.7, Table 14.40.

Figure 10.12: Effect of adjusting for child, family, school and area characteristics upon teacher externalising score (complete case analysis)



For full results, see Appendix 2, Section 14.7, Table 14.41.

Level 1 variables: area, school and family SEP

Among the Level 1 variables, only housing tenure had a large effect in reducing ('explaining') the difference between Whites and Indians and only area deprivation had a large effect in increasing ('unmasking') the difference. The other family SEP factors had only modest effects which, as for area deprivation, were mostly in the direction of increasing somewhat (by 0.05 to 0.15) the difference between Whites and Indians. A sensitivity analysis in B-CAMHS04 found almost identical results for occupational social when using the original NS-SEC classification rather than the 'translated' SOC classification. I therefore use the SOC social class measure in all subsequent analyses.

The only Level 1 variable showing inconsistent findings between parents and teachers was Indian ethnic density. Adjusting for this reduced the Indian advantage on the parent externalising score (regression coefficient -0.10) but increased the advantage on the teacher score (regression coefficient +0.12). This is because whereas parent and teacher externalising symptoms generally declined as ethnic density increased, teacher-reported symptoms were raised in the highest ethnic density category (see Figure 10.10, page 270). Because Indian ethnicity is strongly associated with Indian ethnic density (Table 14.37, Appendix 2), even this relatively small difference led to opposite effects on the regression coefficient.

Level 2 variables: family composition and family stress

The family composition and family stress variables with the largest effects in reducing the Indian advantage were those describing the parent couple, namely family type, parental separation and, more modestly, cohabitation status. Other aspects of family composition had little or no effect, nor did parental mental health or any other stressful life event. This reflects the fact that all these variables either showed little difference between Indian and Whites (e.g. mother's age), showed little effect upon externalising problems (e.g. three-generation family), or were too rare to have much impact at the population level (e.g. family police contact). Adjusting for the poorer family functioning of Indian families did, however, increase the unexplained Indian advantage.

Level 3 variables: child characteristics

Of the variables collected in both datasets, the child characteristics with the largest effects in reducing the Indian advantage were parent-reported learning difficulties and teacher-reported academic difficulties. Adjusting for markers of developmental problems also played a modest role. Other than this, adjusting for neuro-developmental, common or rare disorders had little effect, and this remained the case in sensitivity analyses entering each of the constituent disorders independently. Stressful life events to the child, dyslexia and substance use likewise had only small effects. The Indian advantage did, however, increase after adjustment for general health and, to a lesser extent, internalising problems.

Among the B-CAMHS99 reward and punishment variables, adjusting for the non-physical punishments (sending to room, grounding and shouting) reduced the Indian coefficient; adjusting for giving treats/favourite things or smacking had little effect; and adjusting for praising and ever hitting/shaking the child increased the Indian regression coefficient. That these effects go in different directions reflects the fact that Indian parents reported using non-physical punishments less frequently than Whites, but also reported less praise and a higher frequency of ever hitting or shaking the child.

Among the variables collected in B-CAMHS04 only, adjusting for the parent thinking the child's friends were trouble and social aptitudes reduced the Indian advantage considerably, while adjusting for parent disapproval of friends increased it somewhat. For the social support and other relationships with relatives variables, small sample sizes (<70 Indians for the parent scores and <50 for the teacher scores) meant that the analyses were underpowered and the point estimates unstable. There was, however, no indication that adjusting for these variables was important in explaining the difference between Indians and Whites.

Replication after multiple imputation

I repeated the above 'complete case' analyses in the datasets which I created through multiple imputation. The results were almost identical, with the estimated effect of adjustment upon the Indian coefficient usually being unchanged to two decimal places and never changing by more than 0.03.

Effect of adjusting for child, family, school and area characteristics upon the magnitude of the Indian advantage: preliminary multivariable analyses

Correlation between explanatory variables

The correlation between the continuous and ordered categorical explanatory variables indicated that high collinearity was not a problem in this dataset. The only instance of any explanatory variables having a correlation coefficient of over 0.5 was the correlation of 0.52 between the Ford score and IMD score.

Multivariable analyses

Table 10.5, Table 10.6 and Table 10.7 present the effect upon the White (vs. Indian) regression coefficient of adjusting for the Level 1, 2 and 3 factors respectively. In addition to the full sample analyses, I present nested analyses for 9026 Whites and 329 Indians living in two-parent families.[32] This allowed inclusion of the variables measuring mother's and father's economic activity. It also allows assessment of the importance of different variables after removing any confounding by family type.

In almost all cases in all three Levels, the quadratic and cubic terms which were significant at the 5% level in univariable analyses were non-significant in the multivariable model and/or their removal left the point estimate of the linear term and of other variables in the model almost unchanged. In these cases I excluded the quadratic and cubic terms, in order to facilitate interpretation of the coefficients for these variables (presented in the corresponding full models in Section 14.7, Appendix 2). The Tables' footnotes indicate the few instances in which I did retain the quadratic terms.

Level 1 variables: area, school and family SEP

Table 10.5 and Figure 10.13 present the effect of adjusting for the Level 1 variables upon the regression coefficients for White (vs. Indian) ethnicity – that is, upon the 'unexplained' mean difference between White and Indian externalising scores. As the second line of Table 10.5 shows, the regression coefficient for White (vs. Indian) ethnicity reduced by about a quarter after adjusting for housing tenure (1.08 to 0.73 for parent scores, 1.05 to

[32] The full sample actually contains 9052 Whites and 332 Indians in two-parent families, but 26 Whites and three Indians were in single strata once the analyses were restricted to two-parent families. They were therefore excluded from these analyses.

0.72 for teacher scores). Taken in isolation, this could be interpreted as implying that the higher homeownership of Indians is important in explaining the Indian advantage. Yet as described above, Indians do not uniformly enjoy higher SEP than Whites. Rather housing tenure is an anomalous SEP indicator, with homeownership certainly being more common in Indians but also less socially differentiated (see Figure 10.7, page 260). It is therefore a less sensitive indicator of overall SEP in Indians than Whites. Insofar as homeownership is hypothesised to be *marker* for social advantage (rather than acting upon externalising problems directly), singling out tenure is misleading. I therefore believe that the most meaningful models in Table 10.5 are those on the bottom row which adjust simultaneously for area, school and family SEP variables.

In these models, the adjusted regression coefficient in the full sample returns to close to its unadjusted value (0.97 vs. 1.08 for parent scores; 1.11 vs. 1.05 for teacher scores). In the nested analyses of two-parent families the regression coefficient for White (vs. Indian) ethnicity in fact *increases* after adjustment for the Level 1 variables. This reflects the fact that White two-parent families are over-represented among socio-economically advantaged groups, while in Indians this association is much less strong (Figure 10.7, p.260).

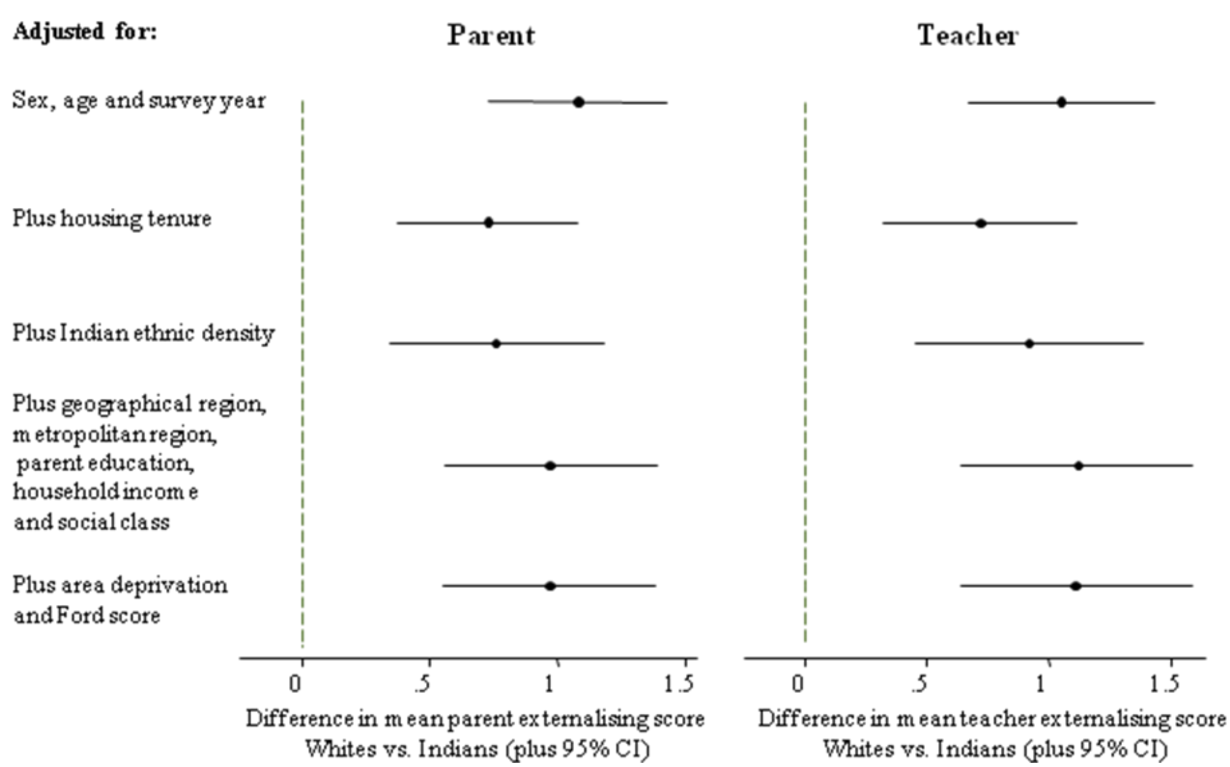
These results therefore indicate that the Indian advantage cannot be explained by ethnic differences in the area, school and family SEP variables. Note that in Table 10.5, and indeed in almost all subsequent analyses, the parent and teacher externalising scores produced very similar findings. This adds considerably to the confidence one can have in the substantive conclusions. As in univariable analyses, the main exception was that adjusting for Indian ethnic density increased the regression coefficient for the teacher scores while leaving it little changed for parent scores (for this reason I present the variable on a separate line). I examine the cause of this discrepancy in more detail in Section 11.1.1 Chapter 11

Table 10.5: Effect of adjustment for all Level 1 variables upon the regression coefficient for White (vs. Indian) ethnicity

	Full sample		Nested analysis: two-parent families	
Adjusted for:	**Parent externalising score (13868 White, 361 Indian)**	**Teacher externalising score (10775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus housing tenure	0.73 (0.37, 1.08)***	0.72 (0.32, 1.11)***	0.52 (0.14, 0.89)**	0.50 (0.06, 0.95)*
Plus Indian ethnic density	0.76 (0.34, 1.18)***	0.92 (0.45, 1.39)***	0.46 (0.05, 0.86)*	0.70 (0.23, 1.17)**
Plus geographical region, metropolitan region, parent education, household income and social class	0.97 (0.56, 1.39)***	1.12 (0.64, 1.59)***	0.69 (0.28, 1.10)**	0.92 (0.46, 1.39)***
Plus area deprivation and Ford score	0.97 (0.55, 1.38)***†	1.11 (0.64, 1.59)***	0.69 (0.28, 1.10)**	0.93 (0.47, 1.40)***
[Plus mother's and father's economic activity: nested analysis]			0.73 (0.32, 1.15)**	0.98 (0.51, 1.44)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Indian ethnic density was entered as a linear plus quadratic term in the teacher analysis. †Full model presented in Table 14.42, Section 14.7, Appendix 2

Figure 10.13: Forest plot of the effect of adjustment for all Level 1 variables upon the regression coefficient for White (vs. Indian) ethnicity (full sample)



Level 2 variables: family composition and family stress

Table 10.6 and Figure 10.14 present the effect of adjusting for the Level 2 variables upon the regression coefficient for the difference between the externalising scores of Indians and Whites. Adjusting for family type and parental divorce reduced the Indian advantage substantially (1.08 to 0.62 for parent scores, and 1.05 to 0.61 for teacher scores). These coefficients increased somewhat, however, after adjusting for other aspects of family composition and other stressful life events. The variable contributing most to this increase was mother's age at the child's birth; when this was entered in addition to family type and parental separation, the regression coefficient increased from 0.62 to 0.72 for parent scores and 0.61 to 0.68 for teacher scores. This reflects the fact that, as described above, younger White mothers were particularly overrepresented in step- and lone parent families. As such, while there was no difference in mother's age between Indians and Whites overall, Indian mothers in two-parent families were younger than their White counterparts. As younger mother's age is associated with higher externalising scores, adjusting for this increases the unexplained difference between Whites and Indians. This also suggests that the detrimental effect of family type in Whites is partly due to confounding by mother's age. Adjusting only for family type may therefore overestimate how much of Indian advantage has been explained.

Finally, adjusting for family functioning increased the regression coefficient for the Indian advantage substantially (from 0.85 to 1.09 for parent scores, and 0.81 to 0.92 for teacher scores), thus bringing it back almost to its the initial level. This reflects the fact that family functioning is poorer in Indian families and therefore acts as a negative confounder when entered into the model.

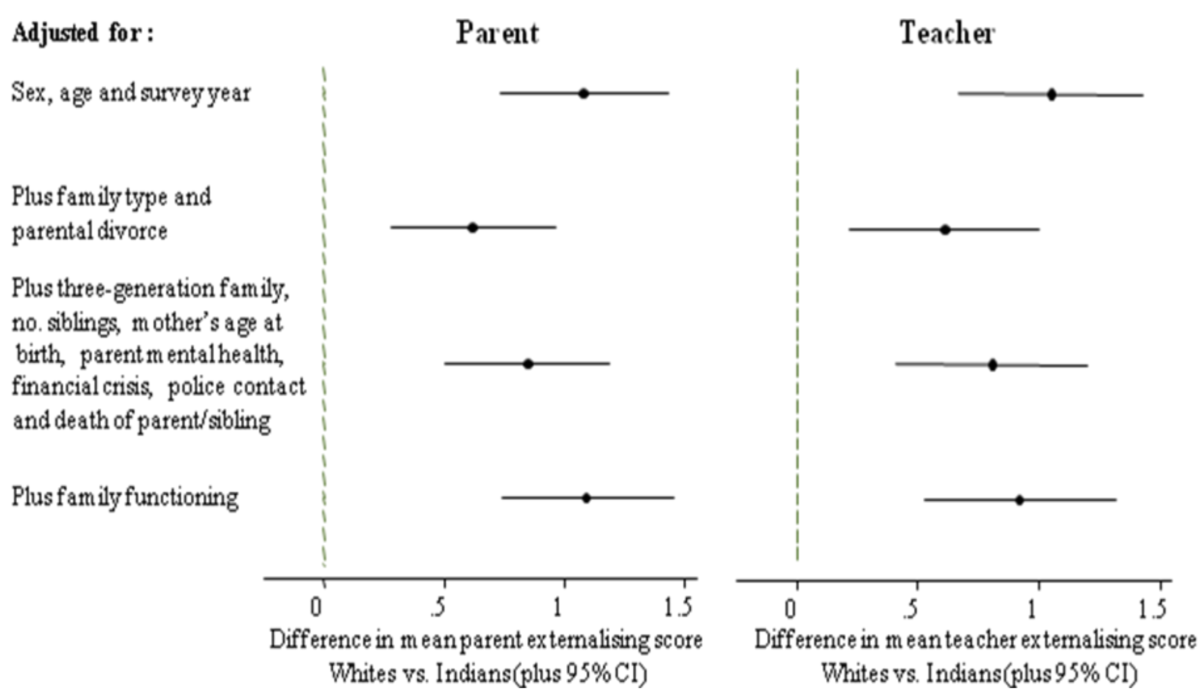
These substantive findings were similar when repeated in two-parent families, and additionally adding parent marital status to the model had little further effect.

Table 10.6: Effect of adjustment for all Level 2 variables upon the regression coefficient for White (vs. Indian) ethnicity

	Full sample		Nested analysis: two-parent families	
Adjusted for:	**Parent externalising score (13868 White, 361 Indian)**	**Teacher externalising score (10775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus family type and parental divorce	0.62 (0.28, 0.96)***	0.61 (0.22, 1.00)***	0.56 (0.19, 0.92)**	0.55 (0.12, 0.99)*
Plus three-generation family, no. co-resident siblings, mother's age at child's birth, parent mental health, family financial crisis, family police contact and death of parent or sibling	0.85 (0.50, 1.19)***	0.81 (0.41, 1.20)***	0.79 (0.42, 1.16)***	0.80 (0.35, 1.24)***
Plus family functioning	1.09 (0.74, 1.45)***†	0.92 (0.53, 1.32)***	1.03 (0.65, 1.42)***	0.89 (0.45, 1.34)***
[Plus parent marital status: nested analysis]			1.00 (0.61, 1.38)***	0.87 (0.43, 1.32)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 14.43, Section 14.7, Appendix 2

Figure 10.14: Forest plot of the effect of adjustment for all Level 2 variables upon the regression coefficient for White (vs. Indian) ethnicity (full sample)



Level 3 variables: child characteristics

Table 10.7 and Figure 10.15 present the effect of adjusting for the child characteristics collected in both B-CAMHS surveys. Adjusting for teacher-reported academic difficulties, parent-reported learning difficulties and developmental problems reduced the regression coefficient for the Indian advantage from 1.08 to 0.77 for parent scores, and 1.05 to 0.76 for teacher scores. Almost all of this effect was attributable to the two academic variables – adjusting only for teacher-reported academic difficulties and parent-reported learning difficulties gave point estimates of 0.80 for parent scores and 0.76 for teacher scores. Additionally adjusting for specific disorders, stressful life events and dyslexia had little further effect. These substantive findings were similar in the nested analyses of two-parent families.

Among the child characteristics, academic abilities therefore seemed the most important factor in explaining the Indian mental health advantage. As described in Chapter 9 (p.240), B-CAMHS99 also administered formal tests of children's reading and spelling. I conducted a sensitivity analysis in B-CAMHS99 to assess the effect of additionally adjusting for these formal tests. In models adjusting only for teacher-reported academic difficulties, parent-reported learning difficulties and parent-reported dyslexia (plus age, gender and survey year) the regression coefficients for White (vs. Indian) ethnicity were 0.58 for parent scores and 0.92 for teachers. These changed little upon additionally adjusting for the child's spelling and reading scores (as continuous variables), becoming 0.51 for parent scores and 0.91 for teachers. This suggests these formal tests do not substantially reduce residual confounding. This is reassuring in that it implies that my inability to adjust for these tests in analyses of both B-CAMHS surveys is not a major limitation and is unlikely to affect my substantive conclusions.

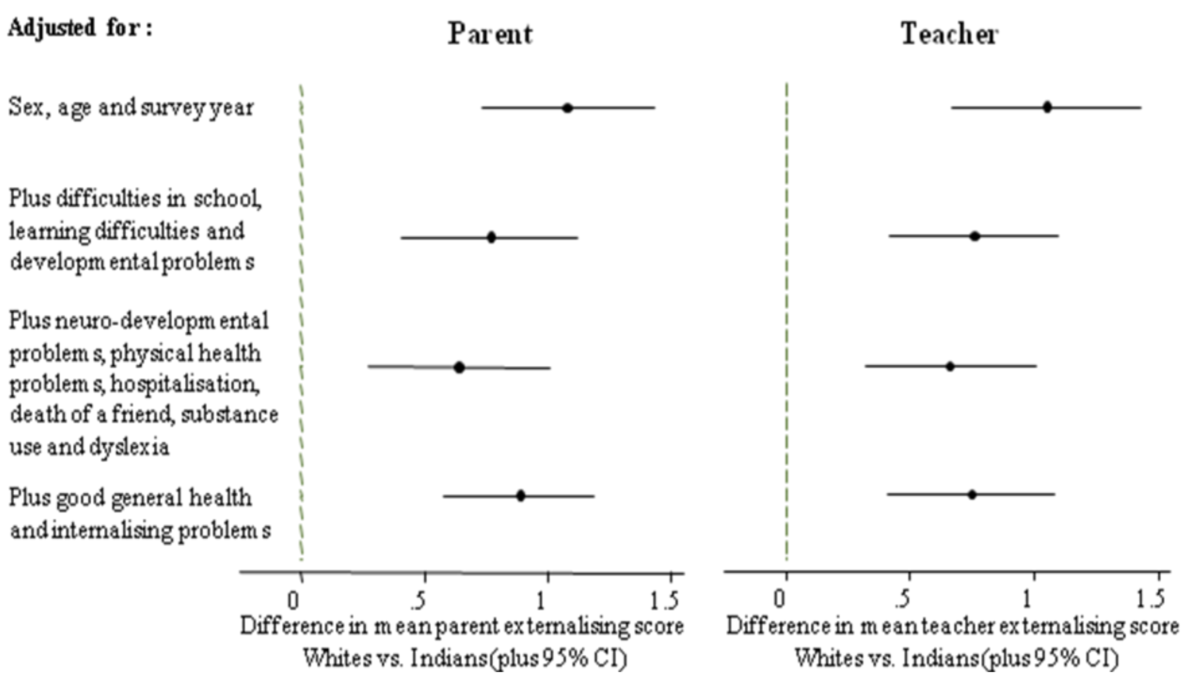
The third line of Table 10.7 adjusts for a large number of child characteristics. Each of these had only small effects in univariable analyses, but always in the direction of reducing the regression coefficient (Figure 10.11). It is therefore unsurprising that their cumulative effect is a modest decrease of the regression coefficient (from 0.77 to 0.64 for parent scores and from 0.76 to 0.66 for teacher scores). Additionally adjusting for parent-reported general health and internalising problems increased the coefficient again (to 0.89 for parent scores and 0.75 for teacher scores). This increase was mostly driven by adjustment for general health. As I discuss in more detail in Chapter 11, the evidence of a reporting bias for this variable means this adjustment for general health may not in fact be appropriate.

Table 10.7: Effect of adjustment for all Level 3 variables upon the regression coefficient for White (vs. Indian) ethnicity

Adjusted for:	Full sample		Nested analysis: two-parent families	
	Parent externalising score (13868 White, 361 Indian)	Teacher externalising score (10775 White, 257 Indian)	Parent externalising score (9026 White, 329 Indian)	Teacher externalising score (7156 White, 237 Indian)
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus academic difficulties, learning difficulties and developmental problems	0.77 (0.41, 1.12)***	0.76 (0.42, 1.10)***	0.50 (0.13, 0.88)**	0.50 (0.14, 0.87)**
Plus neuro-developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, regular smoker, alcohol consumption, ever used drugs and dyslexia	0.64 (0.27, 1.01)***	0.66 (0.32, 1.01)***	0.40 (0.02, 0.78)*	0.45 (0.08, 0.82)*
Plus good general health and internalising problems.	0.89 (0.58, 1.19)***†	0.75 (0.41, 1.08)***	0.67 (0.34, 1.00)***	0.52 (0.16, 0.89)**

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 14.44, Section 14.7, Appendix 2

Figure 10.15: Forest plot of the effect of adjustment for all Level 3 variables upon the regression coefficient for White (vs. Indian) ethnicity (full sample)



Level 3 variables: child characteristics only from B-CAMHS-99 or B-CAMHS04

B-CAMHS99 collected information on parental strategies for rewarding and punishing their children. Table 10.8 presents the effect of adjusting for these in the B-CAMHS99 sample, in addition to the child characteristics collected in both surveys. Adjusting for rewards and punishments reduced the Indian advantage and rendered it only weakly significant (from 0.78 to 0.50 for parent scores, p=0.06; and from 0.76 to 0.58 for teacher scores, p=0.03). Most of this decrease was due to the effect of adjusting for the non-physical punishments sending to room, grounding and shouting. Indeed, when adjusting only for these three punishments, the difference between Indians and Whites reduced to only 0.32 for the parent score and became highly non-significant (p=0.23).

Table 10.8: Effect of additionally adjusting for parent-reported variables on rewards and punishments (B-CAMHS99 only)

Adjusted for:	Full sample	
	Parent externalising score (7872 White, 194 Indian)	Teacher externalising score (6298 White, 145 Indian)
Sex, age and survey year	0.80 (0.28, 1.32)**	1.07 (0.52, 1.61)***
Plus academic difficulties, learning difficulties and developmental problems; neuro-developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, regular smoker, alcohol consumption, ever used drugs, dyslexia; general health and internalising problems	0.78 (0.30, 1.25)**	0.76 (0.30, 1.22)**
Plus send to room; grounding; shouting	0.32 (-0.21, 0.86) [p=0.23]	0.56 (0.05, 1.06)*
Plus treats; favourite things; smacking	0.33 (-0.20, 0.86) [p=0.22]	0.56 (0.05, 1.07)*
Plus praise; ever hits/shakes child	0.50 (-0.02, 1.02) [p=0.06]	0.58 (0.07, 1.09)*

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression.

B-CAMHS04 collected three parent-reported measures of relations with peers. Table 10.9 presents the effect of adjusting for these in the B-CAMHS04 sample, in addition to the child characteristics collected in both B-CAMHS surveys. Social aptitudes and the parent thinking the child's friends were trouble had substantial further effects upon the regression coefficient. Specifically, additionally adjusting for these two variables decreased the regression coefficient for the Indian advantage, from 1.00 to 0.66 for parent scores, and from 0.71 to 0.53 for teacher scores, although this increased again somewhat after additionally adjusting for parent disapproval of friends.

Table 10.9: Effect of additionally adjusting for parent-reported variables on peer relations (B-CAMHS04 only)

Adjusted for:	Full sample	
	Parent externalising score (5996 White, 167 Indian)	Teacher externalising score (4477 White, 112 Indian)
Sex, age and survey year	1.42 (1.04, 1.80)***	1.03 (0.51, 1.55)***
Plus academic difficulties, learning difficulties and developmental problems; neuro-developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, regular smoker, alcohol consumption, ever used drugs, dyslexia; general health and internalising problems	1.00 (0.65, 1.34)***	0.71 (0.21, 1.22)**
Plus social aptitudes and parent thinks friends are trouble	0.66 (0.22, 1.09)**	0.52 (0.00, 1.05) [p=0.05]
Plus parent disapproves of friends	0.73 (0.29, 1.16)**	0.59 (0.06, 1.12)*

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Social aptitudes was entered as a linear plus quadratic term

Table 10.10 presents the effect of adjusting for the child-reported variables on relations with relatives in the B-CAMHS04 subpopulation of 11-16 year olds. These had almost no additional effect, changing the parent regression coefficient from 0.73 to 0.72 and the teacher coefficient from 0.81 to 0.80. This lack of effect is unsurprising given the very similar distribution in Indians and White for social support and number of close relatives, and given that helping relatives showed no association with externalising problems.

Table 10.10: Effect of additionally adjusting for child-reported variables on relations with relatives (11-16 year olds from B-CAMHS04 only)

Adjusted for:	Full sample	
	Parent externalising score (3034 White, 87 Indian)	Teacher externalising score (2105 White, 51 Indian)
Sex, age and survey year	1.18 (0.65, 1.71)***	1.14 (0.43, 1.86)**
Plus academic difficulties, learning difficulties and developmental problems; neuro-developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, regular smoker, alcohol consumption, ever used drugs, dyslexia; general health and internalising problems	0.73 (0.19, 1.27)**	0.81 (0.09, 1.52)*
Plus social support, number of close relatives inside and outside the home, helping relatives.	0.72 (0.15, 1.29)*	0.80 (0.04, 1.56)*

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression.

Summary of findings and implications

These findings therefore indicated that there was little overall effect of adjusting for the Level 1 variables relating to area characteristics, school characteristics and family SEP. Among the Level 2 variables, a higher prevalence of two-parent families and a lower prevalence of parental separation seemed to explain a substantial fraction of the Indian advantage. This effect remained, albeit attenuated, after adjusting for other aspects of family composition such as mother's age at the child's birth. Adjusting for the apparently poorer family functioning of Indian families had the opposite effect, however, increasing the unexplained Indian advantage. Among the Level 3 child characteristics, the Indian mental health advantage seemed to be partly explained by Indian children having fewer academic difficulties. The Indian regression coefficient was also decreased by adjusting for non-physical punishments, social aptitudes and the parent thinking the child's friends are trouble, but was increased by adjusting for general health.

These findings suggest that living more often in two-parent families, doing better in school, receiving fewer non-physical punishments, having better social aptitudes and having fewer friends who are trouble may be important contributing factors to the Indian mental health advantage. The findings are, however, only preliminary. This is because when seeking to explain the Indian advantage, there may be some explanatory variables which should not be entered into the model or which should not be entered as main effects. For example, I have already demonstrated that the general health variable shows inconsistencies with other health measures which are suggestive of a reporting bias between Indians and Whites. Other possible reasons why the above all-variable, main effects analyses could be inappropriate include interactions with ethnicity and/or reverse causality. I address these issues in the next Chapter, and use the findings to inform further multivariable analyses.

Chapter 11 Aim three, part three: Understanding the Indian advantage – interactions, reverse causality and final multivariable analyses

The first two sections of this Chapter address two issues which the previous Chapter ignored: interactions and reverse causality. This then informs further, more sophisticated multivariable analyses in the third section, investigating which explanatory variables seem most important in explaining the Indian mental health advantage.

11.1 Further considerations in explaining the Indian advantage: interactions and reverse causality

11.1.1 Interactions between ethnicity and selected explanatory variables

<u>Rationale for analyses</u>

The univariable and multivariable analyses in Chapter 10 modelled all potential explanatory variables as main effects, and recorded how adjusting for these variables affected the magnitude of the difference in externalising scores between Indians and Whites. This is of interest in identifying mediators or confounders which seem important in 'explaining' the observed Indian advantage. Also of potential interest is to look for interactions between explanatory variables and ethnicity. This can distinguish the possibility that the Indian advantage applies equally across all levels of an explanatory variable (i.e. no interaction) from the possibility that the Indian advantage is greater for some levels than for others (an interaction). Such interactions may generate rich aetiological insights into why particular characteristics affect externalising problems. Identifying such interactions is also important in order to fit appropriate multivariable models.

The previous section reported that there was no evidence of an interaction between ethnicity and the child's age or gender. In this Section, I have decided *a priori* to investigate interactions between Indian ethnicity and three further types of characteristics: Indian ethnic density, socio-economic disadvantage and family structure.

As discussed in Chapter 9 (p.229), Indian ethnic density is the only variable in my conceptual model for which an interaction with Indian ethnicity is central to the mechanism whereby it is hypothesised to affect child mental health. This follows from the hypothesis that it is the density of the individual's *own* ethnic group that is particularly important (see Chapter 2, p.39). This effect is usually hypothesised to be protective, such that living in an area in which one's own group makes up a higher fraction of the population promotes good mental health [154, 158]. Protective effects may, however, be masked by a concentration of minority ethnic groups in socio-economically deprived areas. It is therefore vital to adjust for area deprivation when investigating ethnic density effects [156] – a conclusion reinforced by my own previous demonstration that area deprivation was highest in areas of very low or very high Indian ethnic density (Figure 10.5, p.257).

My decision to test for interactions with socio-economic disadvantage was motivated by Maugham's demonstration in B-CAMHS99 that the marked White gradient in reading ability by SEP[33] was not observed in Indians [538]. I therefore decided to examine whether a similar interaction existed for externalising problems as an outcome. I used five indicators of socio-economic disadvantage: area deprivation, parent's education, household income, housing tenure and occupational social class.

Lastly, I investigated family structure in terms of both family type (two-parent, step- or lone parent families) and three-generation family status. I selected family type because I believe that Indians and Whites may differ in the circumstances of parental separation – for example, the causes of separation or the degree of ongoing support from the extended family. This could plausibly lead to differences in the implications of family type for externalising problems. Certainly interactions between family type and mental health have been observed for adults in Britain, with lone parenthood being a risk factor for common

[33] Operationalised as a composite measure based upon low occupational social class, low income and low parental education.

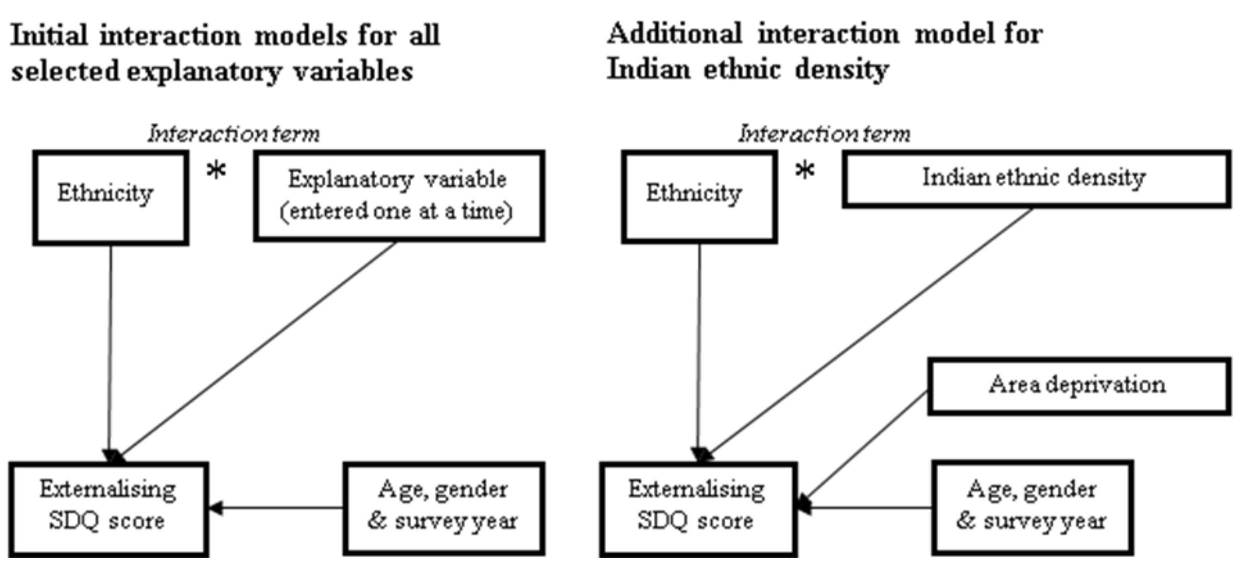
mental health problems in White and South Asian women but not in Black Caribbean women [243]. In addition, I decided to investigate interactions between ethnicity and living in a three generation family. A previous small study of Indian and Pakistani children in London found some evidence of fewer child mental health problems in households containing a grandparent [539]. The study did not include Whites, but it seems plausible that the circumstances surrounding three generation families may differ by ethnic group, and therefore so too may the implications for child mental health.

Methods

I fitted linear regression models which included interactions between ethnicity and each of the selected explanatory variables in turn, plus additionally adjusting for age, gender and survey year (Figure 11.1). For Indian ethnic density, I then additionally adjusted for area deprivation because of its importance as a potential confounder.

I treated housing tenure, social class, family type and three-generation family as categorical variables. For parent education and household income I present the results of entering these variables both as linear terms and as categorical variables, focussing on the former because of the greater power which this offers to detect interactions. I entered Indian ethnic density and area deprivation as a linear terms or, if the quadratic term was significant in either ethnic group modelled separately ($p<0.05$), linear plus quadratic terms.

Figure 11.1: Models testing for interactions between ethnicity and selected explanatory variables

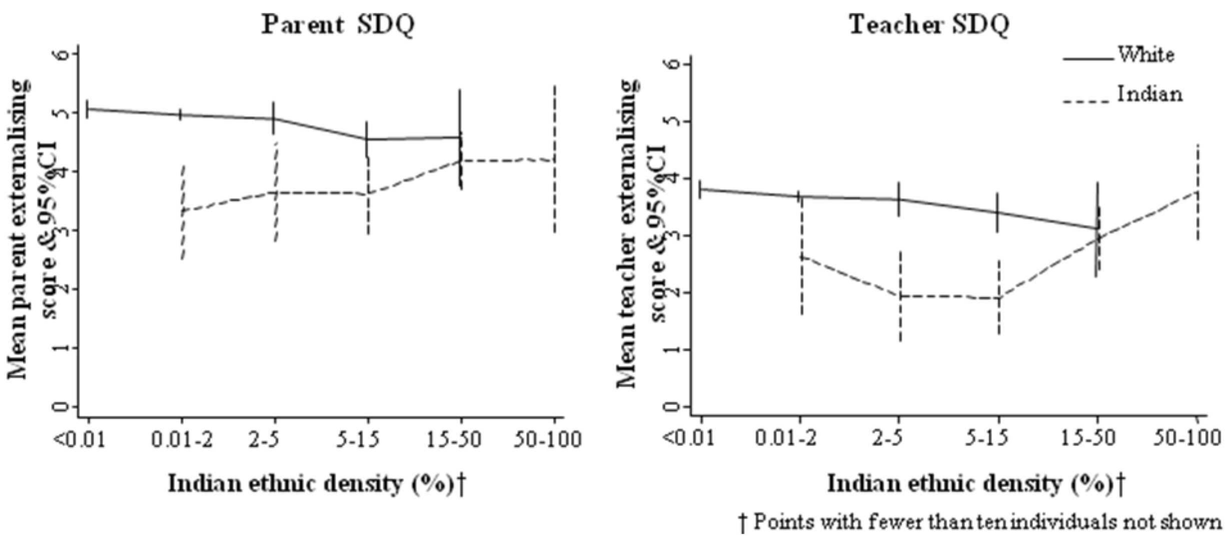


Results

Indian ethnic density

There was weak evidence of an interaction in with Indian ethnic density for both the parent score (p=0.04, linear term) and the teacher score (p=0.03, linear plus quadratic term). In both informants, White externalising problems tended to decrease at higher Indian ethnic densities while Indian externalising problems tended to increase (Figure 11.2). These opposite trends explain the U-shaped relationship between Indian ethnic density teacher scores when Indians and Whites are analysed together (see Figure 10.10 in Chapter 10); at low Indian ethnic density the White values dominate the group average, while the reverse is true at high Indian ethnic density.

Figure 11.2: Parent and teacher externalising scores for Indians and Whites by Indian ethnic density



Indian ethnic density therefore showed weak evidence of an interaction with child's ethnicity in analyses adjusting only for age, sex and survey year. Surprisingly, however, higher Indian ethnic density seemed to be specifically protective to *White* children. More importantly, after adjusting for area deprivation, there was no longer evidence of an interaction with Indian ethnic density for either the parent (p=0.13) or teacher score (p=0.12). Moreover, the trend was still for higher Indian ethnic density to be an advantage in Whites but a disadvantage in Indians. The (non-significant) effect was therefore still not in the expected direction, and also not in the direction necessary to explain the Indian mental health advantage. Taken together, these findings do not support the existence of ethnic density effects in this sample.

Socio-economic disadvantage

All measures of socio-economic disadvantage showed evidence of an interaction with Indian ethnicity such that the deprivation gradient of externalising problems was less marked in Indians than in Whites (Table 11.1 and Figure 11.3). Moreover, not only was the gradient flatter (and in some cases almost flat) in Indians, but the absolute values at the most advantaged end were almost the same. In other words, there was little or no Indian mental health advantage among the most socio-economically advantaged families – instead the advantage was largely confined to less privileged groups.

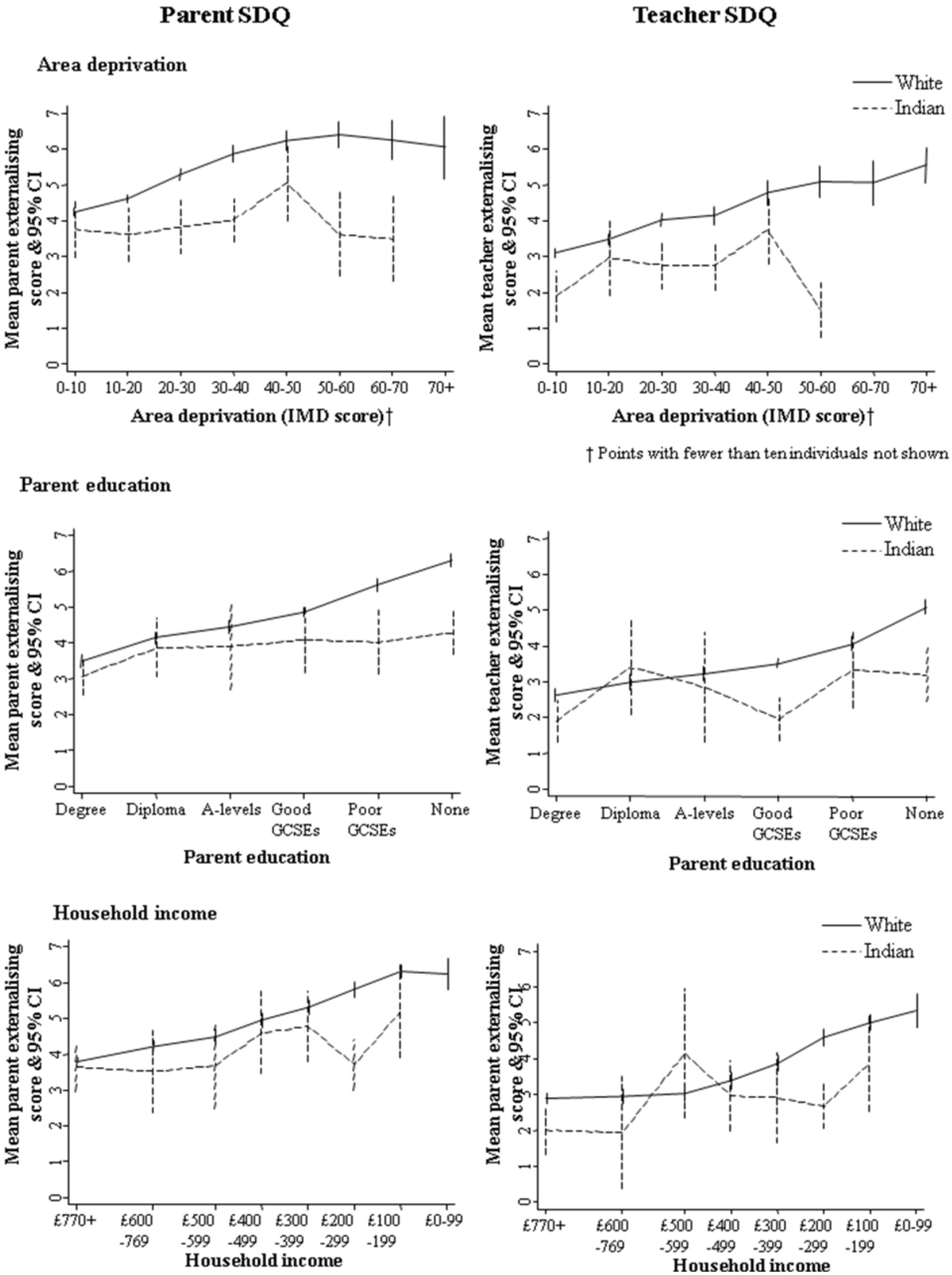
That *all* SEP/area deprivation indicators showed this pattern is very important. If the interaction were seen on just one or two indicators then this might imply that it resulted from the different pattern of inter-relationship between the SEP indicators in Indians and Whites. For example, home ownership is less socially differentiated in Indians than in Whites (Figure 10.6, p.259) and it would therefore be unsurprising if housing tenure were less strongly associated with mental health in Indians. In fact, however, the interaction is also seen for parent education, income and social class which show similar degrees of social differentiation in Indians and Whites (Figure 10.6, p.259). This consistency across all indicators therefore implies that the observed SEP interactions cannot readily be explained as an artefact, but rather may reflect a genuine flattening of the socio-economic gradient in Indians.

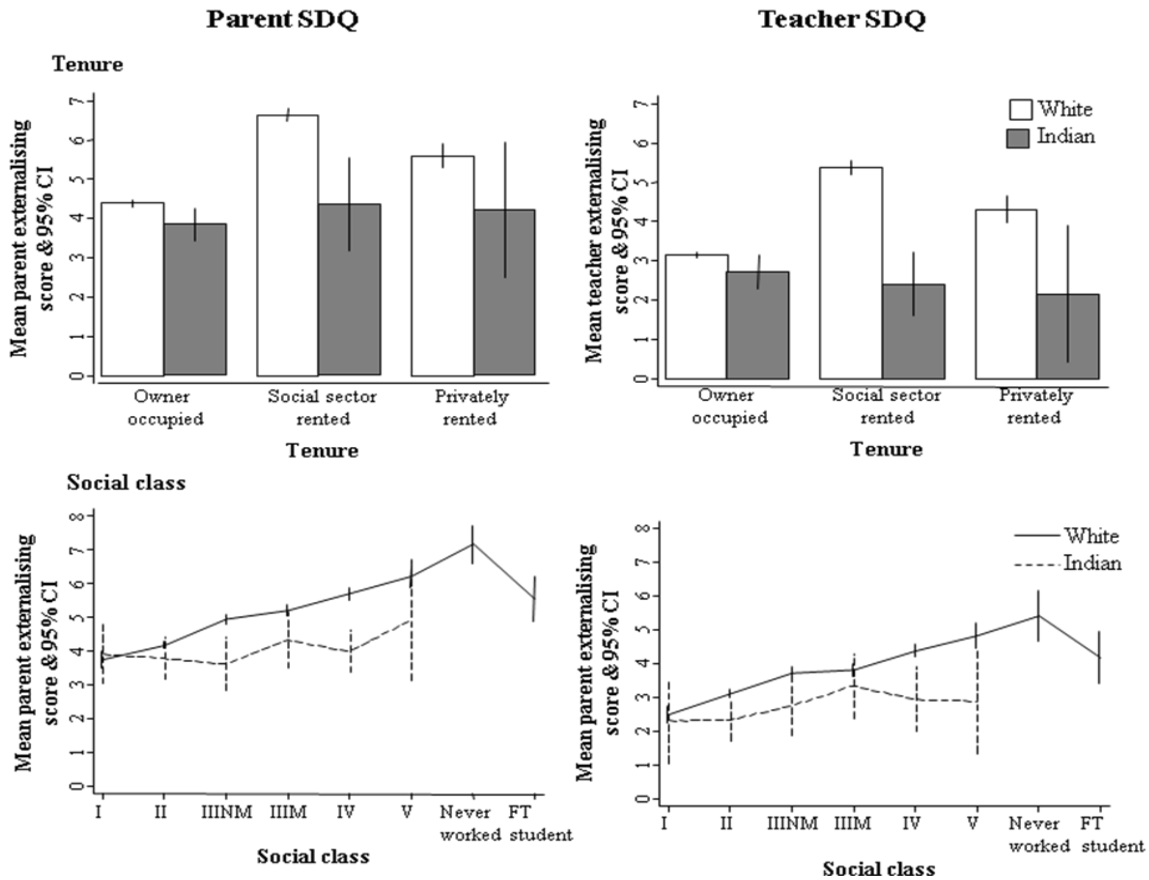
Table 11.1: P-values for interactions between ethnicity and socio-economic disadvantage for the parent and teacher externalising scores

P-value for interaction between ethnicity and:	Parents	Teachers
Area deprivation	0.03	0.008
Parent's education	<0.001 [0.006 if categorical]	0.02 [0.06 if categorical]
Household income	<0.001 [0.002 if categorical]	0.06 [0.02 if categorical]
Housing tenure	0.02	<0.001
Social class	0.01	0.49

Area deprivation, parent education and household income were entered as linear terms, housing tenure and social class categorical. All models were linear regression models with interaction terms between ethnicity and each explanatory variable in question, adjusting for age, gender and survey year.

Figure 11.3: Parent and teacher externalising scores for Indians and Whites by socio-economic disadvantage





Family structure

There was no evidence of an interaction between ethnicity and family type for the parent scores (p=0.17) or the teacher scores (p=0.59).

For living in a three generation family, there was strong evidence of an interaction by teacher report (p=0.001). This was such that there was little difference in Indians between children who lived with a grandparent and those who did not (mean externalising score 2.17 vs. 2.76), but a large difference in Whites (4.82 vs. 3.68). There was, however, no evidence of this interaction for the parent externalising score (p=0.92). There was likewise no evidence of this interaction for the externalising DAWBA diagnosis (p=0.93) or the child-reported externalising score (p=0.23). The interaction for the teacher score therefore seems likely to be a chance finding. This echoes my conclusion in Chapter 10 that the main effect association between living in a three generation family and teacher score was not seen in other informants and therefore likely to be due to chance.

Implications for subsequent analyses

To summarise, these analyses provided no convincing evidence that the density of one's own ethnic group is an important determinant of externalising problems in this sample. Moreover, the trend was for higher Indian ethnic density to be protective for White children. This is the opposite direction to that hypothesised, and also the opposite direction to that necessary to explain the observed Indian mental health advantage. I therefore conclude that ethnic density effects do not seem to be operating in this sample, and certainly cannot explain why Indian children in B-CAMHS have better mental health. This conclusion undermines my rationale for including Indian ethnic density as an explanatory variable, which was predicated on the hypothesis that ethnic density effects might be operating. By contrast, I know of no reason to believe that Indian ethnic density should predict child mental health as a main effect irrespective of the child's ethnicity, and would suspect that any observed relationship in fact reflected the influence of an unmeasured area-level confounder. I therefore intend treat Indian ethnic density with caution and to present subsequent multivariable analyses which do not include it.

By contrast, evidence of an interaction between Indian and SEP/area disadvantage was far more convincing. Across all five indicators, the Indian advantage appeared to be particularly large in socio-economically disadvantaged groups with little or no ethnic difference in the most advantaged group. I believe this finding is of substantial interest, and considerable potential importance. It indicates that the correct question may not be 'Why do Indians (as a whole) have a mental health advantage?' but rather 'What protects Indian children against the negative effects of low SEP?' or alternatively 'What is creating a strong socio-economic gradient in Whites but not in Indians?'. Later in this Chapter I return to this issue by examining how far this interaction is explained by the characteristics measured in B-CAMHS and presenting analyses stratified by SEP. I then discuss this important finding further in Chapter 12.

Finally, there was no convincing evidence of an interaction between ethnicity and family type of three generation families. I therefore continue to enter these variables as main effects into subsequent models.

11.1.2 Reverse causality between externalising problems and child and family factors

The primary purpose of this Chapter is to investigate further which child, family, school and area characteristics are important in explaining the Indian mental health advantage. Yet it would not be correct to 'explain' the Indian advantage with reference to a characteristic which was actually an outcome of their better mental health.

This represents an important problem because some reverse causality is highly plausible between externalising problems and many family stress and child variables (Table 11.2). Some reverse causality may also apply to parental separation, to comparatively non-specific common disorders like migraines, and to illnesses requiring hospitalisation. And, while not plausible as a true cause, the presence of externalising problems might draw attention to dyslexia or developmental problems.

I therefore used the B-CAMHS follow-up surveys to explore how far reverse causality may exist for these child and family covariates, and thereby to inform my multivariable analyses of the Indian advantage.

Table 11.2: Plausibility of reverse causality for family stress and child characteristics

Domain	Reverse causality highly plausible	Reverse causality possible	Reverse causality implausible
Family stress	• Parent mental health • Family functioning†	• Parent separation	• Financial crisis • Family police contact • Death of parent or sibling
Child characteristics	• General health • All substance use variables • Teacher-reported academic difficulties • Parent-reported learning difficulties† • Internalising problems • All rewards variables† • All punishments variables† • All relations with peers variables†† • All relations with relatives variables†††	• Developmental problems† • Common physical disorder† • Serious illness leading to hospitalisation • Dyslexia†	• Death of friend • Neuro-developmental disorder† • Rare physical disorder†

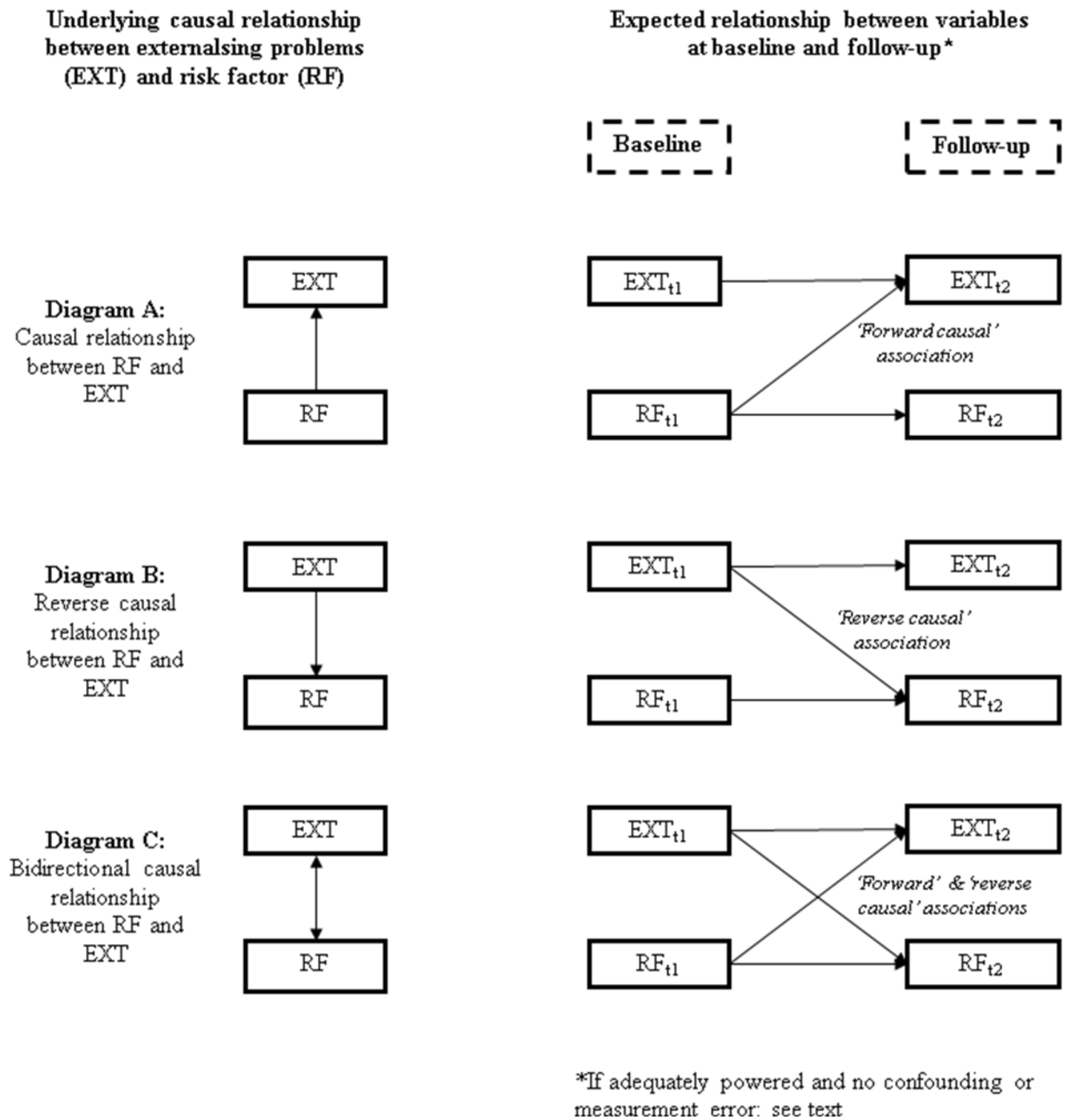
Substance use variables: smoking, alcohol consumption, drug use. Rewards variables: praises, treats, favourite things. Punishments: sending to room, grounding, shouting, smacking, ever hits/shakes. Relations with peers variables: parent thinks friends trouble, parent disapproves of friends, social aptitudes scale. Relations with relatives: social support, number of close relatives in the home, number of close relatives outside the home, how often child helps relatives.
†Follow-up data from B-CAMHS99 only. ††Follow-up data from B-CAMHS04 only. †††Follow-up data from B-CAMHS04, but not available to me.

Conceptual issues

To explore the magnitude and direction of causal effects one would ideally use experimental designs. Failing that, the observational data best suited to the task would involve frequent repeated measurement in the same child of all variables of interest. This would allow greater precision in identifying the timing in changes of one variable relative to another. Measures across multiple time-points could also reduce measurement error by allowing one to use techniques such as growth modelling to extract latent trajectories.

The B-CAMHS surveys included only a single main follow-up at three years, and are therefore not well suited to examining causal directions. Nevertheless, the follow-up data does provide some limited purchase on the issue which is not possible in cross-sectional analyses. In general, longitudinal analyses are useful because if a putative risk factor is truly a cause of externalising problems then changes in the risk factor should predict future changes in externalising problems. For example, the onset (or deterioration) of a risk factor would be expected to predict increasing externalising problems. Conversely if the putative risk factor were in fact an outcome then it would not be expected to predict future externalising problems (after adjustment for current externalising problems). Rather one would expect current externalising problems to predict future values of the risk factor. And if the relationship were bidirectional, one would expect each factor to predict the other (see Figure 11.4).

Figure 11.4: Expected longitudinal associations between putative risk factors and externalising problems, given different underlying causal relationships



Of course even if a putative risk factor does predict future externalising problems (a 'forward causal' association), a genuine causal effect is not the only explanation. In observational studies, alternative explanations include unmeasured confounders; measurement error resulting in incomplete adjustment for baseline externalising problems; or antecedents of externalising problems having preceded the risk factor but at 'sub-clinical' levels not detected at baseline (see Figure 11.5).[34] This applies equally to 'reverse causal' associations, given that risk factors may also be measured with error or be insidious in their onset. This includes even apparently clear-cut events like parental divorce or parental death, which may in fact represent the culmination of a long period of family tension or deteriorating parent health.

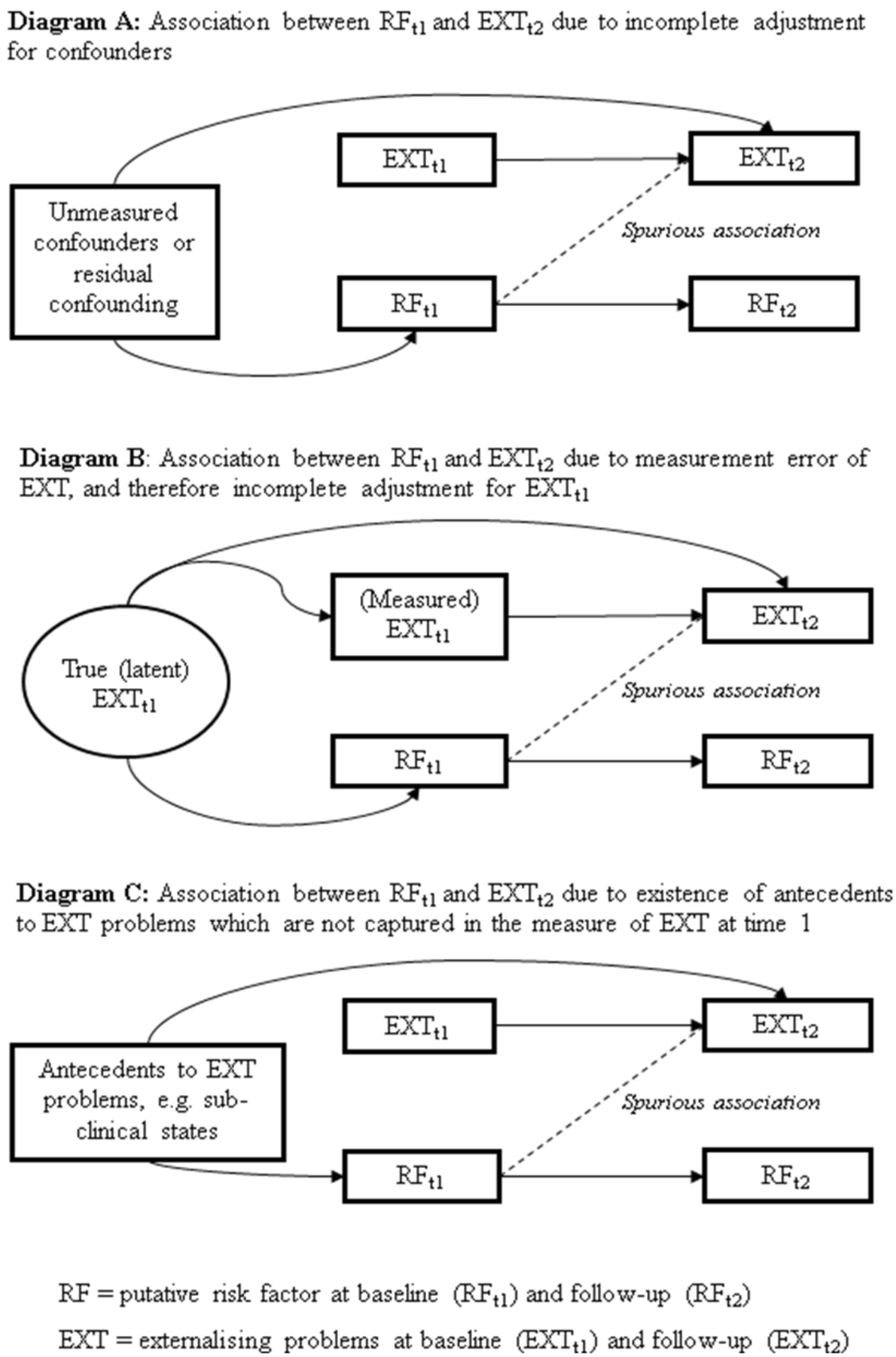
The observation that a risk factor at baseline predicts externalising problems at follow-up is thus *consistent* with causality, but not the only explanation. Still, if no future predictive relationship exists this may mean the relationship is *not* causal. I will therefore consider excluding from my analyses any variable whose relationship with externalising problems takes the form of Diagram B, Figure 11.4. Yet even here, caution is needed – while the absence of association is certainly consistent with non-causality, it is also consistent with causality which is confined to a specific critical window in the past; with underpowered analyses; or with real but transient effects which might no longer be apparent after three years.

In addition to identifying potential reverse causal relationships (Figure 11.4, Diagram B), it would also be of interest to estimate the relative strength of causal effects in bidirectional relationships (Figure 11.4, Diagram C). It is tempting to do so by comparing the magnitude of the effect sizes on the two cross-lagged arrows, using a comparable metric such as the standardised regression coefficients. Yet although potentially informative, considerable caution is needed when applying this approach. This is because measurement error in independent variables tends to reduce the regression coefficient in a regression model (regression dilution bias). By contrast, measurement error in dependent variables does not affect the point estimate of the regression coefficient (although, by increasing the 'unexplained' error, it does widen the confidence intervals). As such, 'reverse causal'

[34] This final possibility is, however, less relevant for dimensional measures like the SDQ, instead applying particularly to binary measures of disorder.

regression coefficients could be larger than ‘forward causal’ coefficients simply because externalising problems were measured with less error than the risk factor in question.

Figure 11.5: Non-causal explanations for apparent causal associations



Methods

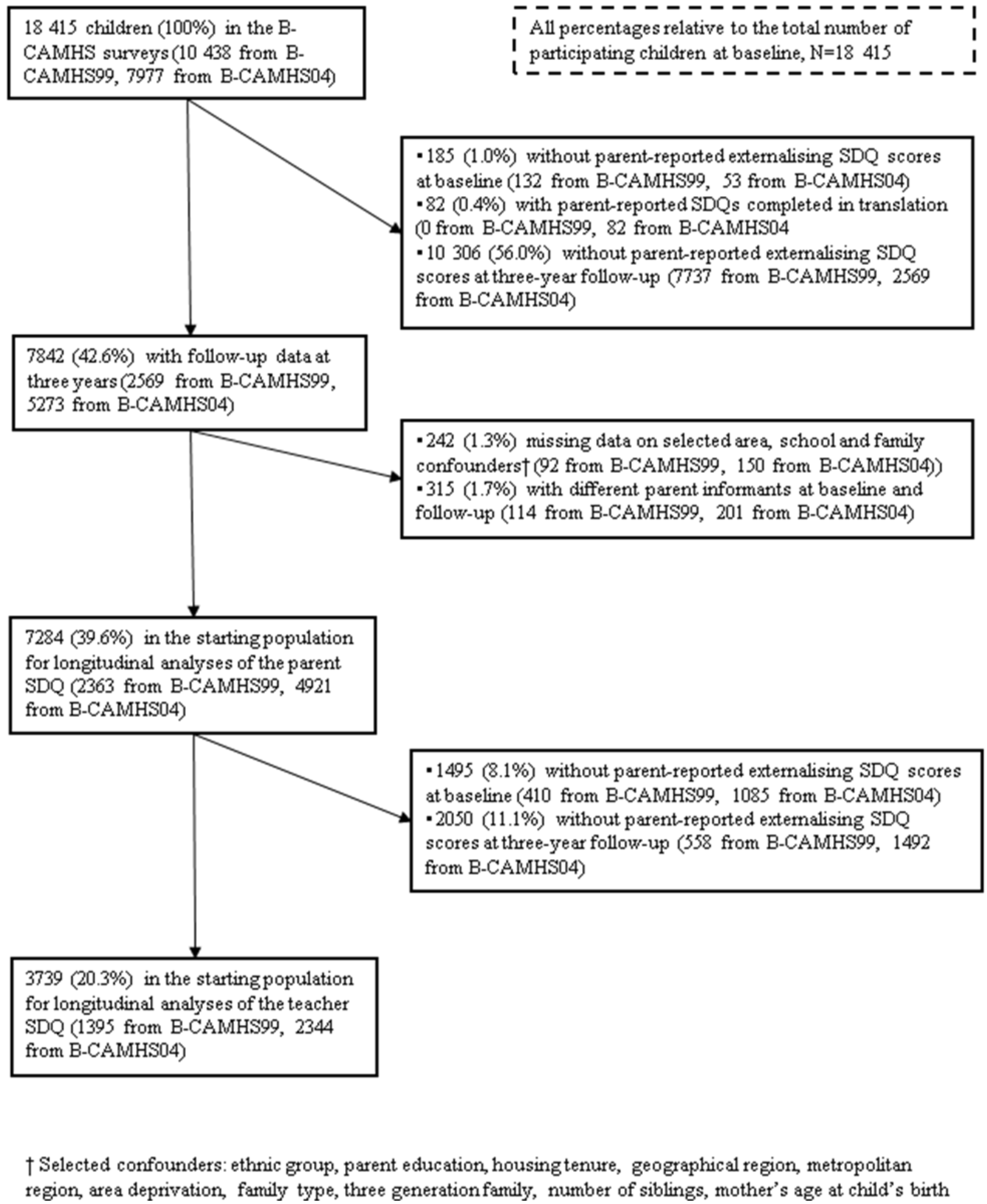
Starting population for causality analyses

I had no reason to hypothesise that causal relationships would vary across the countries of Britain or between ethnic groups. I therefore decided to use the whole B-CAMHS sample for these analyses, rather than only Indian and White children from England. This increased the sample size by about 25%, an important consideration given that insufficient power may prevent detection of genuine causal associations. As a sensitivity analysis, however, I repeated these analyses after restricting the sample to the Indian and White children from England, as used in Chapter 10 (see Figure 10.1).

As shown in Figure 11.6, 7842 individuals (42.6% of the total B-CAMHS sample) had baseline and follow-up data from non-translated parent SDQs. To create a stable subset for my subsequent analyses, I excluded 240 additional children who were missing data on one or more of selected family, area or school variables confounders. These confounders were ethnic group (White, Black Caribbean, Black African, Indian, Pakistani, Bangladeshi, Chinese or 'Other'); parent income; housing tenure; geographical region; metropolitan area; area deprivation; family type; three generation family; number of co-resident siblings; and mother's age at child's birth.[35] I further excluded the 315 individuals with a different informant at baseline and follow-up, in order to allow meaningful interpretation of the parent mental health variable. This left 7284 children (39.6% of the total B-CAMHS population) with parent SDQs and 3739 (20.3%) children with teacher SDQs.

[35] I did not adjust for mother's economic activity, father's economic activity, or parent cohabitation as these variables were only collected for subsets of families. Ethnic density was not available for Scotland and Ford score was not available for Scotland or Wales. Finally, I did not adjust for household income and social class because these had substantial missing data (7.1% for income, 3.1% for social class) and because parent education and housing tenure were already included as SEP indicators. Sensitivity analyses in children for whom all these variables were available indicated that including these additional potential confounders had virtually no effect upon the results, and none upon my substantive conclusions.

Figure 11.6: Starting population for analyses of causal directions



Data availability follow-up and descriptive analyses

Three-year follow-up data from at least one B-CAMHS dataset was available to me for almost all the child and family stress variables. The only exceptions were the child-reported variables about relations with relatives (social support, number of close relatives and helping relatives), which I was unable to obtain from ONS. For most variables, the follow-up assessment was identical to the baseline assessment. The only exceptions were the stressful life events; whereas the baseline survey asked if these had 'ever' happened, the follow-up survey asked if they had happened 'in the last three years'.

The variables listed previously in Table 11.2 (p.295) are those which I use in my substantive analyses of the Indian advantage. I alsoassessed teacher-reported internalising SDQ score, to see if it replicated findings for the parent-reported internalising score.

Descriptive analyses, unadjusted regression and adjusted regression analyses

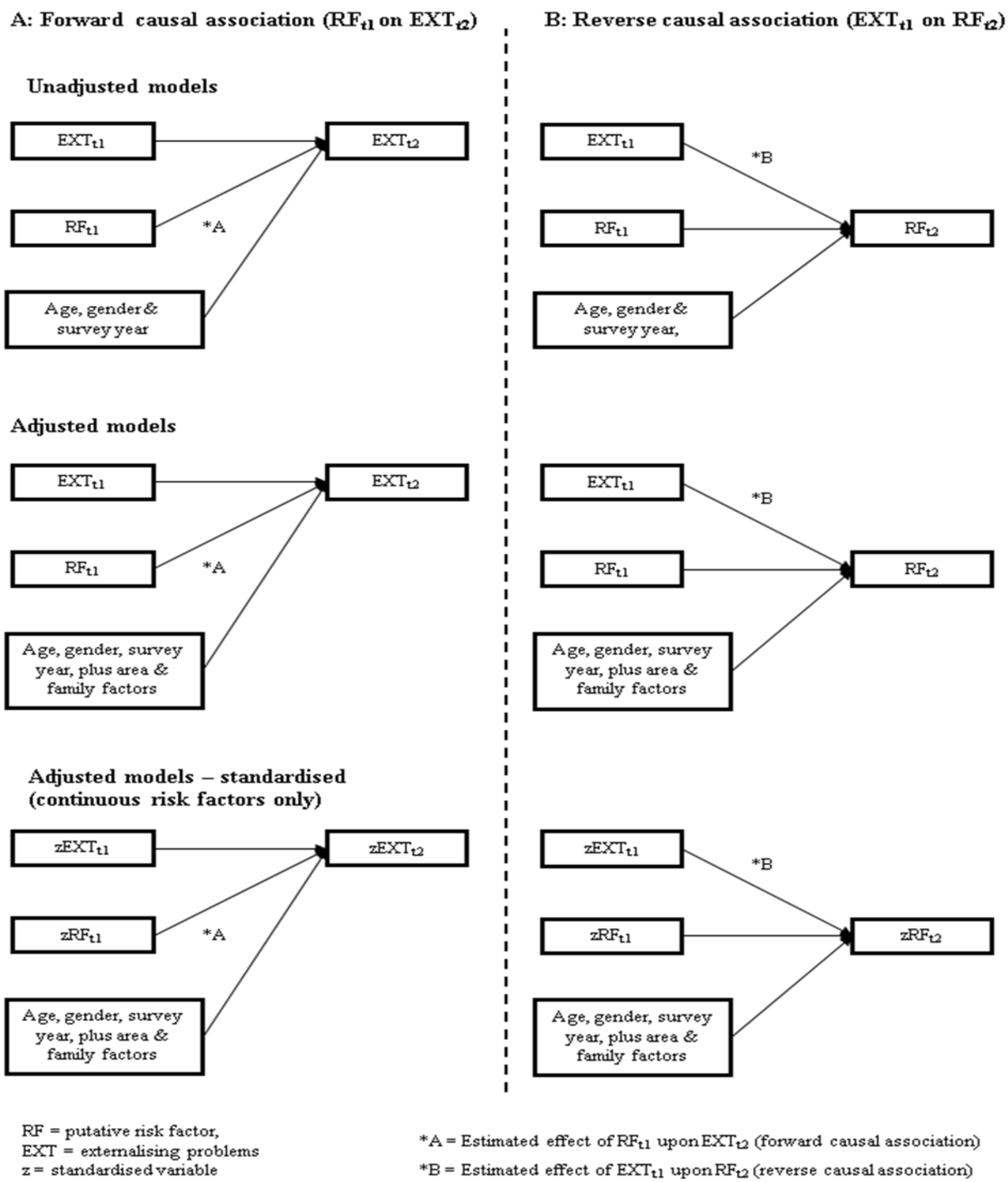
As an initial descriptive analysis, I cross-tabulated externalising scores at follow-up against externalising scores at baseline, and calculated Spearman's correlation coefficients between the two time points. I repeated this for all putative risk factors.

I then fitted a series of unadjusted and adjusted regression models (Figure 11.7):

- **Unadjusted models:** Two regression analyses, one with follow-up externalising scores as the outcome and the other with the follow-up risk factor as the outcome. Both models included baseline externalising and risk factor scores as explanatory variables, plus age, gender and survey year. I used logistic regression for binary risk factors; ordered logistic regression for ordered categorical risk factors; and linear regression for continuous risk factors.
- **Adjusted models**: As in the unadjusted model, but also adjusting for the selected family, area and school confounders. I adjusted for ethnicity, parent education, housing tenure, geographical region, metropolitan region, area deprivation (as country-specific quartiles), family type, three-generation family and number of co-resident siblings as categorical variables; and mother's age at child's birth as a continuous variable.
- **Adjusted models – standardised**: For continuous risk factors, I repeated the adjusted model after standardising the risk factor and externalising scores (i.e.

subtracting the mean and then dividing by the standard deviation, to give transformed scores with mean=0 and standard deviation=1). I standardised the baseline and follow-up scores separately. These standardised models therefore show the number of standard deviations change in the dependent variable per standard deviation change in independent variable.

Figure 11.7: Sequence of models fitted to explore the causal directions



To summarise information from across the many risk factors, I categorised the strength of each forward causal/reverse casual association as follows:

- No evidence of association ($p>0.10$)
- Weak/marginal evidence ($0.01\leq p<0.10$)
- Strong evidence: ($0.001\leq p<0.01$)
- Very strong evidence ($p<0.001$).

Cross-tabulating these gave a table in the form of Table 11.3, with each risk factor occupying a particular square. My primary motivation in conducting these analyses was to identify variables with reverse causal relationships with externalising problems (Figure 11.4, Diagram B). I was therefore particularly interested in variables occupying the bottom left cells (circled), and considered these to be plausible instances of reverse causality.

Table 11.3: Grid for preliminary assessment of likely causal directions

		A: Strength of evidence of forward causal association (RF_{t1} on EXT_{t2})			
		None ($p>0.10$)	**Weak/marginal ($0.01\leq p<0.10$)**	**Strong ($0.001\leq p<0.01$)**	**Very strong ($p<0.001$)**
B: Strength of evidence of reverse causal association (EXT_{t1} on RF_{t2})	**None ($p>0.10$)**	No relation or low power	No relation or low power	Causal	Causal
	Weak/marginal ($0.01\leq p<0.10$)	No relation or low power	No relation or low power	Causal	Causal
	Strong ($0.001\leq p<0.01$)	Reverse causal	Reverse causal	Bidirectional	Bidirectional
	Very strong ($p<0.001$)	Reverse causal	Reverse causal	Bidirectional	Bidirectional

Variables occupying circled cells are of particular interest, in suggesting possible instances of reverse causality.

For the continuous variables, I also tested whether there was evidence of a difference between the magnitude of the forward causal (zRF_{t1} upon $zEXT_{t2}$) and reverse causal ($zEXT_{t1}$ upon zRF_{t2}) standardised regression coefficients. I did this using Z test statistics, calculated as the difference between the coefficients divided by the estimated standard error of the difference (see Appendix 1, p.380). In interpreting these findings, I took into consideration the probable greater measurement error of scales which have been less thoroughly validated than the externalising SDQ subscale and/or which have substantially fewer measurement points.

Results

Table 11.4 describes the distribution of follow-up parent externalising scores for each baseline score; Section 14.8.1 Appendix 2, contains equivalent tables for teachers and all putative risk factors. As Figure 11.8 shows, there was a near-linear relationship between baseline externalising scores and mean follow-up scores. The Spearman's correlation coefficients between the baseline and follow-up scores were 0.69 for parents and 0.53 for teachers (Table 11.5). For the putative risk factors these coefficients varied considerably. For example, they were very low (≤0.11) for the stressful life events; moderate (0.25-0.50) for parent mental health, substance use, most health measures and most rewards/punishment variables; and high (>0.50) for family functioning, academic abilities and neuro-developmental disorders. The corresponding Pearson's coefficients for the continuous variables were very similar to the Spearman's coefficients, varying by ≤ 0.04.

Table 11.4: Distribution of parent externalising scores by externalising score at baseline

Parent externalising score (SDQ points)		**Follow-up**										
		0-1	**2-3**	**4-5**	**6-7**	**8-9**	**10-11**	**12-13**	**14-15**	**16-17**	**18-20**	**All children**
Base line	**0-1**	977	438	165	51	16	6	1	0	0	0	1654
	2-3	519	604	352	135	53	7	11	2	0	0	1683
	4-5	194	417	370	266	95	40	7	8	4	0	1401
	6-7	63	191	264	221	129	63	24	6	3	1	965
	8-9	19	56	151	172	143	76	42	12	6	2	679
	10-11	4	21	55	78	87	76	41	20	10	3	395
	12-13	2	7	24	37	52	59	40	41	11	3	276
	14-15	0	5	4	11	9	33	22	19	18	4	125
	16-17	0	1	4	4	5	10	12	13	15	11	75
	18-20	0	1	0	1	3	2	1	6	10	7	31
	All children	1778	1741	1389	976	592	372	201	127	77	31	7284

Spearman's correlation coefficient 0.69

Figure 11.8: Mean follow-up externalising score by externalising score at baseline

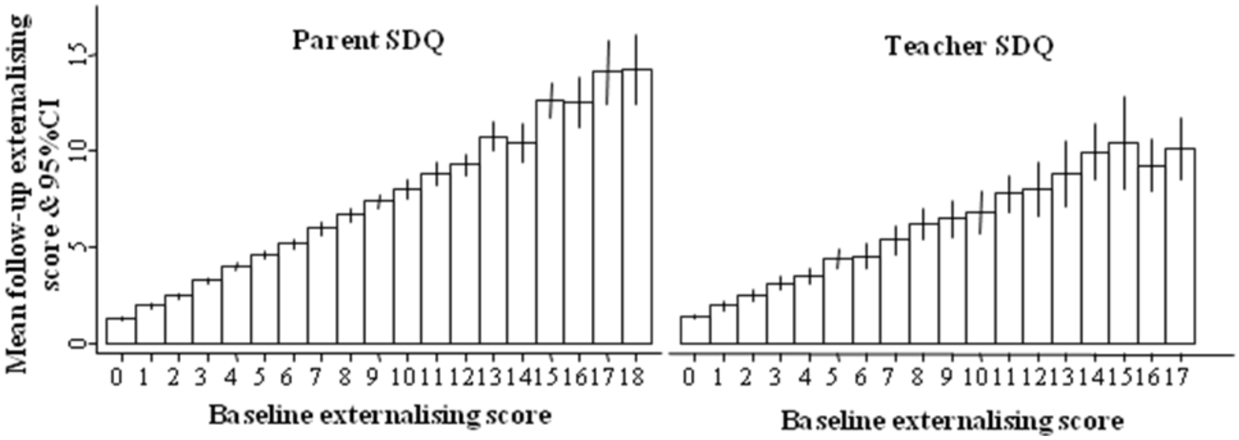


Table 11.5: Correlation between baseline and follow-up variables

Variable	N	Spearman's correlation
Parent externalising scores	7284	0.69
Teacher externalising scores	3739	0.53
Parental separation	7261	0.10
Financial crisis	7263	0.11
Family police contact	7262	0.07
Death of parent or sibling	7264	0.03
Parent mental health	7235	0.37
Family functioning	2343	0.55
Good general health	7280	0.40
Neuro-developmental disorder	2363	0.74
Developmental problems	2363	0.53
Common physical disorder	2363	0.51
Rare physical disorder	2363	0.42
Serious illness leading to hospitalisation	7263	0.05
Death of friend	7263	0.06
Regular smoker	6284	0.43
Alcohol consumption	6281	0.42
Ever used drugs	6279	0.40
Teacher-reported academic difficulties	2596	0.65
Learning difficulty	2363	0.62
Dyslexia	2363	0.62
Parent-reported internalising problems	7281	0.55
Teacher-reported internalising problems	3768	0.33
Reward: praise	1750	0.29
Reward: treats	1746	0.30
Reward: favourite things	1734	0.26
Punish: send to room	1748	0.44
Punish: grounding	1748	0.45
Punish: shouting	1750	0.42
Punish: smacking	1751	0.44
Punish: ever hit or shake	1751	0.21
Parent disapproval of friends	4783	0.21
Parent thinks friends are trouble	4772	0.26
Social aptitudes score	4910	0.62

Regression analyses: Identifying variables with reverse causal relationships with externalising problems

Table 11.6 and Table 11.7 summarise the strength of the evidence for the forward and reverse causal associations between the child and family stress variables and parent/teacher externalising scores. These grids are based on the adjusted models; full results of all unadjusted and adjusted models are presented in Section 14.8.2 Appendix 2. In general, the results using the parent and teacher externalising scores were similar, although the smaller sample size (≈50%) for the teacher analyses meant the results were less highly powered and the significance levels reduced. The substantive findings were also very similar or identical in analyses restricted only to White and Indian children from England.

As outlined above, my primary interest in these analyses is to identify variables which do *not* predict externalising problems, but which are strongly predicted by them (Diagram B, Figure 11.4). There were six variables which showed this reverse causal relationship with both parent and teacher externalising scores. These were smoking, alcohol consumption, drug use, punishment through shouting, parent disapproval of friends and the parent thinking the child's friends were trouble. The three reward variables also showed this reverse causal relationship with parent externalising scores, and the weaker evidence of reverse causal associations in teachers may be due to particularly small sample sizes (N≤1750 because these were only collected in B-CAMHS99). I therefore consider all nine variables to be plausible outcomes, not causes, of externalising problems.

The evidence for a reverse causal association was also stronger than that for a forward causal association for grounding, smacking or sending the child to their room. Moreover, evidence for a forward causal effect was confined to the highest punishment level (see Figure 11.9). This is consistent with the fact that, uniquely among the punishment variables, the most severe punishment category (ever hits or shakes the child) showed stronger evidence of a forward causal than a reverse causal association. I therefore intend to retain grounding, smacking and sending the child to their room, but to recode them into binary 'frequently'/'not frequently' variables.

Finally, family police contact showed a reverse causal relationship with the parent externalising score. Remembering that this variable is about police contact regarding a family member *other* than the child, I do not believe genuine reverse causality is plausible here. Instead, unmeasured confounding (e.g. shared genetic risk) or chance are more likely explanations, with the lack of replication on the teacher externalising score making chance particularly plausible. This variable therefore exemplifies the need for caution when drawing conclusions throughout this section: these analyses identify patterns of association which are *consistent* with reverse causality, but this may not be the only explanation.

Table 11.6: Summary of strength of forward and reverse causal associations between child and family stress risk factors and the parent externalising score

		A: Strength of evidence of forward causal association (RF_{t1} on parent EXT_{t2})			
		None ($p>0.10$)	**Weak/marginal ($0.01\leq p<0.10$)**	**Strong ($0.001\leq p<0.01$)**	**Very strong ($p<0.001$)**
B: Strength of evidence of reverse causal association (parent EXT_{t1} on RF_{T2})	**None ($p>0.10$)**	• Neuro-developmental disorder	• Rare physical problems • Punish: ever hit or shake	• Common physical problems	
	Weak/ marginal ($0.01\leq p<0.10$)	• Death of parent or sibling • Child hospitalisation • Death of a friend			
	Strong ($0.001\leq p<0.01$)	• Family police contact • Rewards: treats • Rewards: favourite things	• Financial crisis		
	Very strong ($p<0.001$)	• Smoking • Alcohol consumption • Drug use • Rewards: praise • Punish: shouting	• Punish: grounding† • Parent disapproval of friends • Parent thinks friends are trouble	• Parental separation • Developmental problems • Dyslexia • Punish: send to room† • Punish: smack†	• Parent mental health • Family functioning • Good general health • Teacher-reported academic difficulties • Learning difficulty • Internalising problems (parent) • Internalising problems (teacher) • Social aptitudes

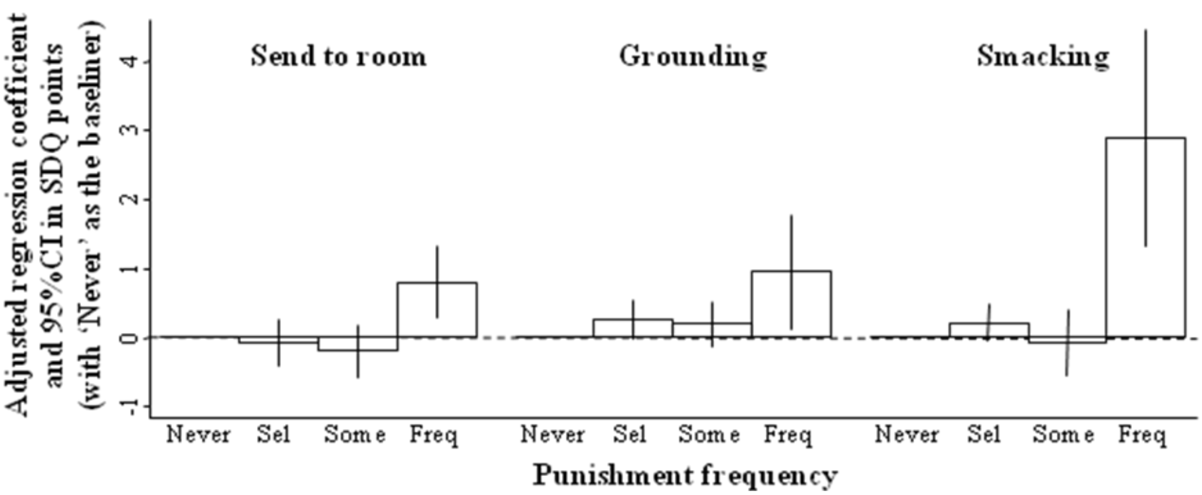
†Evidence of causal relationship only for top category. For full results of models, see Table 14.46 and Table 14.47, Section 14.8.1, Appendix 2

Table 11.7: Summary of strength of forward and reverse causal associations between child and family stress risk factors and the teacher externalising score

		A: Strength of evidence of forward causal association (RF_{t1} on teacher EXT_{t2})			
		None (p>0.10)	**Weak/marginal (0.01≤p<0.10)**	**Strong (0.001≤p<0.01)**	**Very strong (p<0.001)**
B: Strength of evidence of reverse causal association (teacher EXT_{t1} on RF_{t2})	**None (p>0.10)**	• Neuro-developmental disorder • Rare physical problems • Rewards: praise • Rewards: favourite things	• Family police contact • Common physical problems • Punish: ever hit or shake • Family functioning	• Good general health	• Parental separation • Dyslexia
	Weak/ marginal (0.01≤p< 0.10)	• Financial crisis • Death of a friend • Rewards: treats • Parent mental health	• Death of parent or sibling		
	Strong (0.001≤p <0.01)	• Child hospitalisation • Punish: shouting	• Punish: send to room† • Punish: smack		
	Very strong (p<0.001)	• Smoking • Drug use • Parent thinks friends are trouble	• Alcohol consumption • Internalising problems (parent) • Internalising problems (teacher) • Parent disapproval of friends	• Punish: grounding† • Social aptitudes	• Developmental problems • Teacher-reported academic difficulties • Learning difficulty

†Evidence of causal relationship only for top category. For full results of models, see Table 14.48 and Table 14.49, Section 14.8.1, Appendix 2

Figure 11.9: Adjusted effect of punishment at baseline upon the follow-up parent externalising score ('forward causal' association)



'Sel'=seldom, 'Some'=sometimes, 'Freq'=frequent. For underlying regression models, see Table 14.46, Section 14.8.1 Appendix 2

Most of the remaining variables showed bidirectional relationships (Diagram C, Figure 11.4) with the parent and teacher externalising scores. For the continuous variables, Table 11.8 and Table 11.9 present the standardised regression coefficients from the adjusted models. Different degrees of measurement error mean that caution is necessary when comparing the magnitude of these coefficients. Nevertheless, it was notable that the reverse causal regression coefficients were 2.5-4 times larger than the forward causal coefficients for parent-reported internalising problems, teacher-reported internalising problems and social aptitudes. This was replicated across the parent and teacher externalising scores and the difference in the coefficients was always highly significant ($p \leq 0.002$). This is particularly striking because these scales have at least as great a range as the externalising scores (0-20 for the internalising and externalising scores, 0-40 for the social aptitudes scale). This argues against the larger reverse coefficients simply reflecting cruder measurement of internalising problems or social aptitudes.

In addition, parent mental health showed a substantially stronger reverse causal than forward causal association with parent externalising scores (p-value for difference <0.001). The magnitudes of both correlation coefficients were substantially smaller for teacher scores,[36] and the difference between them was not statistically significant. Yet while there was no evidence of an effect of parent mental health on teacher externalising scores ($p=0.13$) there was some evidence of an effect in the reverse direction ($p=0.03$). Together, this provides evidence that parent mental health may also primarily be an outcome rather than a cause of externalising problems.

By contrast, for the remaining three continuous variables there was not convincing evidence that one causal direction predominated over another. General health showed evidence of a larger reverse causal association in parents ($p=0.004$) but this was not replicated for teachers. Moreover, the fact that general health was measured on only a five-point scale means that greater measurement error is a very plausible explanation for the larger reverse causal association observed with parent externalising scores. Teacher-reported academic difficulties and family functioning showed only weak evidence ($p=0.03$)

[36] This may indicate that shared rater bias explains some of the association between parent-reported externalising scores and parent mental health. Alternatively/additionally it may indicate genuine situation-specific effects, e.g. it is child externalising problems *at home* which particularly affect parent mental health.

of a difference in the magnitude of coefficients for parent externalising scores and no evidence of difference for teacher scores.

Table 11.8: Standardised regression coefficients from adjusted models for continuous risk factors and the parent externalising score

	N	A: RF_{t1} on parent EXT_{t2} (forward causal association); standardised regression coefficient and 95%CI	B: Parent EXT_{t1} on RF_{t2} (reverse causal association); standardised regression coefficient and 95%CI	P-value for difference between regression coefficients
Parent mental health	7235	0.038 (0.020, 0.055)***	0.118 (0.093, 0.144)***	<0.001
Family functioning	7243 (A)/ 2343 (B)	0.041 (0.024, 0.058)***	0.085 (0.048, 0.122)***	0.03
Good general health	7280	-0.032 (-0.050, -0.014)***	-0.073 (-0.095, -0.051)***	0.004
Teacher-reported academic difficulties	2596	0.066 (0.033, 0.099)***	0.116 (0.087, 0.144)***	0.03
Parent-reported internalising problems	7281	0.046 (0.028, 0.064)***	0.129 (0.105, 0.153)***	<0.001
Teacher-reported internalising problems	3768	0.058 (0.033, 0.082)***	0.136 (0.102, 0.170)***	<0.001
Social aptitudes score	4910	-0.064 (-0.085, -0.042)***	-0.178 (-0.202, -0.153)***	<0.001

*p<0.05, **p<0.01, ***p<0.001.

Table 11.9: Standardised regression coefficients from adjusted models for continuous risk factors and the teacher externalising score

	N	A: RF_{t1} on teacher EXT_{t2} (forward causal association); standardised regression coefficient and 95%CI	B: Teacher EXT_{t1} on RF_{t2} (reverse causal association); standardised regression coefficient and 95%CI	P-value for difference between regression coefficients
Parent mental health	3176	0.021 (-0.006, 0.048)	0.038 (0.005, 0.072)*	0.44
Family functioning	3728 (A)/ 1389 (B)	0.027 (-0.002, 0.055)*	0.019 (-0.029, 0.068)	0.80
Good general health	3735	-0.042 (-0.071, -0.014)**	-0.026 (-0.059, 0.008)	0.46
Teacher-reported academic difficulties	2537	0.105 (0.069, 0.142)***	0.078 (0.041, 0.114)***	0.28
Parent-reported internalising problems	3739	0.029 (0.000, 0.059) [p=0.05]	0.098 (0.065, 0.132)***	0.002
Teacher-reported internalising problems	3735	-0.026 (-0.055, 0.003) [p=0.08]	0.109 (0.070, 0.148)***	<0.001
Social aptitudes score	2334	-0.055 (-0.092, -0.019)**	-0.138 (-0.175, -0.101)***	0.002

*p<0.05, **p<0.01, ***p<0.001.

Summary and implications for subsequent analyses

To summarise, these results indicate that several child and family stress variables may have a reverse causal relationship with externalising problems. Specifically, this is true of parent mental health; smoking, alcohol consumption and drug use; parent- and teacher-reported internalising problems; rewards through praise, treats, and favourite things; punishment through shouting; the parent disapproving of the child's friends; the parent thinking the child’s friends are trouble; and social aptitudes. Punishment through grounding, smacking or sending the child to their room also showed strong reverse causal associations and forward causal associations which were confined to the top punishment category.

I conclude that all these variables may primarily be outcomes not causes of externalising problems. This conclusion is substantially strengthened by the consistency across similar variables: all three substance use variables; both internalising mental health measures; all three reward measures; all four common punishments at low levels (but, interestingly, *not* at high levels or for the one rare and severe form of punishment); and all three peer relations variables. This conclusion is also strengthened by the replication of these findings across the parent and teacher externalising scores. Finally, it is reassuring that ‘rewards through treats’ and ‘rewards through favourite things’ were among the variables showing a reverse causal relationship. As discussed in Chapter 10 (p.271), the unexpected positive association between these two variables and higher externalising scores seems most likely to reflect parents using rewards/treats to manage difficult children. It is therefore an encouraging indication of the validity of the approach used in this Section that it supports this conclusion.

These findings are important for my PhD because it would not be correct to ‘explain’ the Indian advantage in terms of variables which are in fact outcomes of externalising problems. I therefore intend to repeat the preliminary multivariable analyses presented in Chapter 10 after the excluding variables which showed reverse causal relationships with externalising problems. I will also only include grounding, smacking and sending the child to their room after recoding these into binary ‘frequent’/’not frequent’ variables. Nevertheless, it is important to remember that alternative explanations for these apparently reverse causal relationships cannot be ruled out. These alternative explanations include a lack of power (particularly for the rarer substance use variables) and/or poor measurement (particularly for rewards, punishments and parent disapproval of friends, all of which have only three or four levels). As such, I believe the preliminary multivariable analyses in Chapter 10 are still of value as sensitivity analyses of what happens to the Indian advantage when these problematic variables are included.

11.2 Understanding the Indian mental health advantage: final multivariable analyses

Table 11.10 summarises the variables which my previous analyses suggest should not be entered into multivariable models attempting to explain the Indian advantage. I therefore repeat the multivariable analyses reported in Chapter 10 after excluding these problematic variables.

Table 11.10: Summary of problematic variables excluded from subsequent analyses

Reason for exclusion:	Variables affected	Comment
Reporting bias	• General health	Suggestion that Whites systematically report better general health than Indians
Not appropriate as main effect	• Indian ethnic density	Variable included in conceptual model specifically because of hypothesised interaction with ethnicity (this being central to the mechanism of 'ethnic density effects'). No such interaction was observed, however, nor was the trend in the expected direction.
Predominantly reverse causal relationship	• Parent mental health • Internalising problems • Social aptitudes score	All show forward causal associations with externalising scores, but these are substantially smaller than the reverse causal associations.
Reverse causal relationship	• Smoking • Alcohol consumption • Drug use • Rewards: praise • Reward: treats • Reward: favourite things • Punish: shouting • Parent disapproval of friends • Parent thinks friends are trouble	All show little or no evidence of forward causal associations and strong evidence of reverse causal associations.
Reverse causal relationships except at highest frequency	• Punish: send to room • Punish: grounding • Punish: smacking	All show little or no evidence of a forward causal association except at the highest frequency, and strong evidence of a reverse causal association: recode as binary 'Frequent' vs. 'Not frequent'.

11.2.1 Methods

I first repeated the Level-specific multivariable analyses reported in Chapter 10 after excluding the problematic variables identified in Table 11.10. I then progressed to models including variables from multiple Levels. As in Chapter 10, I first entered variables which had the largest univariable effects in reducing the regression coefficient (i.e. 'explaining' the Indian advantage). I then entered variables with little effect and finally variables which increased the regression coefficient. Among the variables with little effect on the regression coefficient, I first entered the Level 1 variables, then Level 2 and finally Level 3, in line with my hierarchical conceptual model.

As a sensitivity analysis, I repeated the final multi-Level model using alternative externalising outcomes (DAWBA diagnosis for externalising disorder and child-reported externalising score). I also repeated the final model with internalising outcomes (DAWBA diagnosis and the three SDQ scores) in order to verify that the adjusted model did not unmask an unexplained difference between Indians and Whites.

I then investigated whether the SEP/area deprivation interactions with ethnicity observed in Section 11.1.1 persisted in multivariable analyses. I did so by calculating the p-value for these interaction terms and by presenting analyses stratified by SEP for each stage of my multi-Level model fitting.

I finish by presenting in full my final multi-variable models of the child, family, school and area predictors of externalising scores in the B-CAMHS dataset.

11.2.2 Results

Level-specific models

Level 1 variables: area, school and family SEP

The effect on the Indian advantage of adjusting for all the Level 1 variables was presented in Table 10.5, Chapter 10. Table 11.11 shows the equivalent analyses excluding Indian ethnic density. The results were relatively little changed. For example, the fully-adjusted parent regression coefficients changed from 0.97 (previously) to 0.98 (Table 11.11) while the teacher coefficients changed from 1.11 to 0.93. After adjusting for the other Level 1 variables in this fully-adjusted model, there was no evidence ($p>0.05$) of either a main effect or interaction term for Indian ethnic density in either the parent or teacher model. I therefore continue to exclude Indian ethnic density from subsequent multivariable models.

The substantive conclusions from Table 11.11 are therefore identical to those in Chapter 10, namely that the Indian advantage is not explained by differences in their area characteristics, school characteristics, or family SEP profile.

Table 11.11: Effect of adjustment for selected Level 1 variables upon the regression coefficient for White (vs. Indian) ethnicity (excluding: Indian ethnic density)

	Full population		**Nested analysis: two-parent families**	
Adjusted for:	**Parent externalising score (13 868 White, 361 Indian)**	**Teacher externalising score (10 775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus housing tenure	0.73 (0.37, 1.08)***	0.72 (0.32, 1.11)***	0.52 (0.14, 0.89)**	0.50 (0.06, 0.95)*
Plus geographical region, metropolitan region, parent education, household income and social class	0.96 (0.60, 1.31)***	0.91 (0.51, 1.31)***	0.76 (0.37, 1.15)***	0.74 (0.28, 1.20)**
Plus area deprivation and Ford score	0.98 (0.61, 1.34)***†	0.93 (0.53, 1.33)***†	0.78 (0.39, 1.17)***	0.77 (0.31, 1.23)**
[Plus mother's and father's economic activity: nested analysis]			0.83 (0.44, 1.22)***	0.85 (0.39, 1.30)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 14.50 and Table 14.51 (Section 14.8.1, Appendix 2).

Level 2 variables: family composition and family stress

The effect on the Indian advantage of adjusting for all the Level 2 variables was presented in Table 10.6, Chapter 10. Table 11.12 shows the equivalent analyses excluding parent mental health. The results were almost identical; from 1.09 to 1.07 for the full-model parent regression coefficients and from 0.92 to 0.91 for teachers. This similarity is unsurprising given the very similar mental health profile of Indian and White parents. As before, therefore, the main conclusion is that the unexplained Indian advantage decreases after adjusting for Indians' higher prevalence of two-parent families but increases after adjusting for Indians' poorer family functioning.

Table 11.12: Effect of adjustment for selected Level 2 variables upon the regression coefficient for White (vs. Indian) ethnicity (excluding: parent mental health)

	Full population		Nested analysis: two-parent families	
Adjusted for:	**Parent externalising score (13 868 White, 361 Indian)**	**Teacher externalising score (10 775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus family type and parental divorce	0.62 (0.28, 0.96)***	0.61 (0.22, 1.00)***	0.56 (0.19, 0.92)**	0.55 (0.12, 0.99)*
Plus three-generation family, no. co-resident siblings mother's age at child's birth, family financial crisis, family police contact and death of parent or sibling	0.77 (0.42, 1.11)***	0.77 (0.37, 1.17)***	0.73 (0.35, 1.11)***	0.75 (0.30, 1.21)**
Plus family functioning	1.07 (0.71, 1.43)***†	0.91 (0.51, 1.31)***†	1.02 (0.63, 1.41)***	0.88 (0.43, 1.33)***
[Plus parent marital status: nested analysis]			0.98 (0.59, 1.37)***	0.86 (0.41, 1.31)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 14.50 and Table 14.51 (Section 14.8.1, Appendix 2).

Level 3 variables (1): child characteristics collected in both datasets

The effect on the Indian advantage of adjusting for all the Level 3 variables was presented in Table 10.7, Chapter 10. Table 11.13 shows the equivalent analyses excluding general health, smoking, alcohol consumption, drug use and internalising problems.

Adjusting for the three substance use variables had previously decreased the regression coefficient for the difference between Indians and Whites. Excluding these variables from the model therefore resulted in the regression coefficients in the third line of Table 11.13 which were larger than their equivalents in Table 10.7 (0.71 vs. 0.64 previously for parents, 0.76 vs. 0.66 previously for teachers). This was offset, however, by effect of also excluding general health and internalising problems, both of which had previously increased the Indian regression coefficient. The final regression coefficient for parent scores was therefore somewhat smaller in Table 11.13 than previously (0.71 vs. 0.89). This indicates that including general health and internalising problems (particularly the former) may previously have generated a misleadingly large estimate of the unexplained difference between Indians and Whites. In teachers the two effects cancelled each other out giving very similar final point estimates (0.76 in Table 11.13 vs. 0.75 previously).

Thus after excluding the problematic variables, the Level 3 child characteristics in Table 11.13 explained about a quarter of the difference between Indians and Whites. As previously, this was largely driven by the measures of academic abilities.

Table 11.13: Effect of adjustment for selected Level 3 variables upon the regression coefficient for White (vs. Indian) ethnicity (excluding: general health, smoking, alcohol consumption, drug use and internalising problems)

Adjusted for:	Full population		Nested analysis: two-parent families	
	Parent externalising score (13 868 White, 361 Indian)	**Teacher externalising score (10 775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus academic difficulties, learning difficulties and developmental problems	0.77 (0.41, 1.12)***	0.76 (0.42, 1.10)***	0.50 (0.13, 0.88)**	0.50 (0.14, 0.87)**
Plus neuro-developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, dyslexia	0.71 (0.35, 1.06)***†	0.76 (0.42, 1.10)***	0.46 (0.09, 0.83)*	0.53 (0.16, 0.90)**

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 14.50 and Table 14.51 (Section 14.8.1, Appendix 2).

Level 3 variables (2): child characteristics collected only in B-CAMHS99 or B-CAMHS04

The only reward and punishment variables showing any evidence of forward causality upon externalising symptoms were frequently sending the child to their room, frequent grounding, frequent smacking and ever hitting/shaking the child. Frequent sending of the child to their room was less common in Indians (3.6% vs. 8.7% in Whites) as was frequent grounding (2.6% vs. 6.0%). By contrast, frequent smacking of the child was somewhat more common in Indians (1.5% vs. 0.5%) as was ever hitting or shaking the child (7.9% vs. 2.6%). These effects largely cancelled each other out when adjusting for these four punishment variables in addition to all the other child characteristics in Table 11.13. Thus in the B-CAMHS99 subset, the Indian regression coefficient changed from 0.48 to 0.44 for the parent scores, and from 0.77 to 0.70 for teacher scores. As such, while the *types* of punishment seemed to vary somewhat between Indians and Whites, these analyses provided no evidence that Indian children were punished less frequently than Whites or that this explained the Indian advantage.

All of the peer relations variables showed evidence of a pattern of reverse causality with externalising symptoms. As such, my finding in Chapter 10 (Table 10.9) that adjusting for these variables substantially decreased the Indian regression coefficient does not imply that these variables truly explain the Indian advantage. Rather, the fact that Indians had better social aptitudes and fewer friends who were trouble seems more likely to be an *outcome*, not a cause, of their reduced prevalence of externalising problems.

Finally, B-CAMHS04 collected child-reported variables about social support, number of close relatives and helping relatives. Unfortunately, I could not obtain the follow-up data necessary to assess evidence for direction of causality in these variables. Moreover, even cross-sectionally these analyses are very underpowered, containing only 87 Indians with parent externalising scores and 51 with teacher scores. As described in Chapter 10 (Table 10.10), univariable analyses provided no evidence that Indians and Whites differed for social support or number of close relatives, and there was no association between helping relatives and externalising problems. Moreover, in the subset of 11-16 year olds from B-CAMHS04, adjusting for these variables in addition to the other child characteristics in Table 11.13 left the parent regression coefficient unchanged (0.83) and changed the teacher coefficient from 0.84 to 0.85. There was therefore no evidence that these relations with relatives variables were important explaining the Indian mental health advantage.

Multi-Level models

In summary, the substantive findings of the multivariable analyses were similar after excluding the problematic variables. The Indian advantage coefficient decreased substantially after adjusting for two-parent families and for academic abilities; increased after adjusting for family functioning; and changed relatively little upon the inclusion of all other variables. This similarity in substantive conclusions is reassuring given the limitations to the methods used to identify problematic variables.

These Level-specific findings guided the sequence in which I fitted multi-Level models, as reported in Table 11.14 and Figure 10.13. Adjusting for academic abilities and learning difficulties decreased the regression coefficient for the Indian advantage by about 20%. Additionally adjusting for family type and parental separation decreased it further still to about half of its initial value (from 1.08 to 0.51 for parent scores, and from 1.05 to 0.51 for teacher scores). Even after this adjustment, however, there remained strong evidence ($p<0.004$) in both informants that Indians were advantaged relative to Whites. Additionally adjusting for area, school and family SEP increased this coefficient somewhat; adding other family composition and family stress variables (other than family functioning) had little effect; and adding other child characteristics decreased the coefficient somewhat for parents and left it unchanged for teachers. All these separate effects were modest (≤0.10 SDQ points), however, with an overall change from 0.51 to 0.54 for parents and 0.51 to 0.62 for teachers. Only adding family functioning had a large effect, increasing the regression coefficient from 0.54 to 0.75 for parent scores and from 0.62 to 0.70 for teacher scores. In the nested analysis of two-parent families, there was little additional effect of adjusting for mother's and father's economic activity or parent marital status.

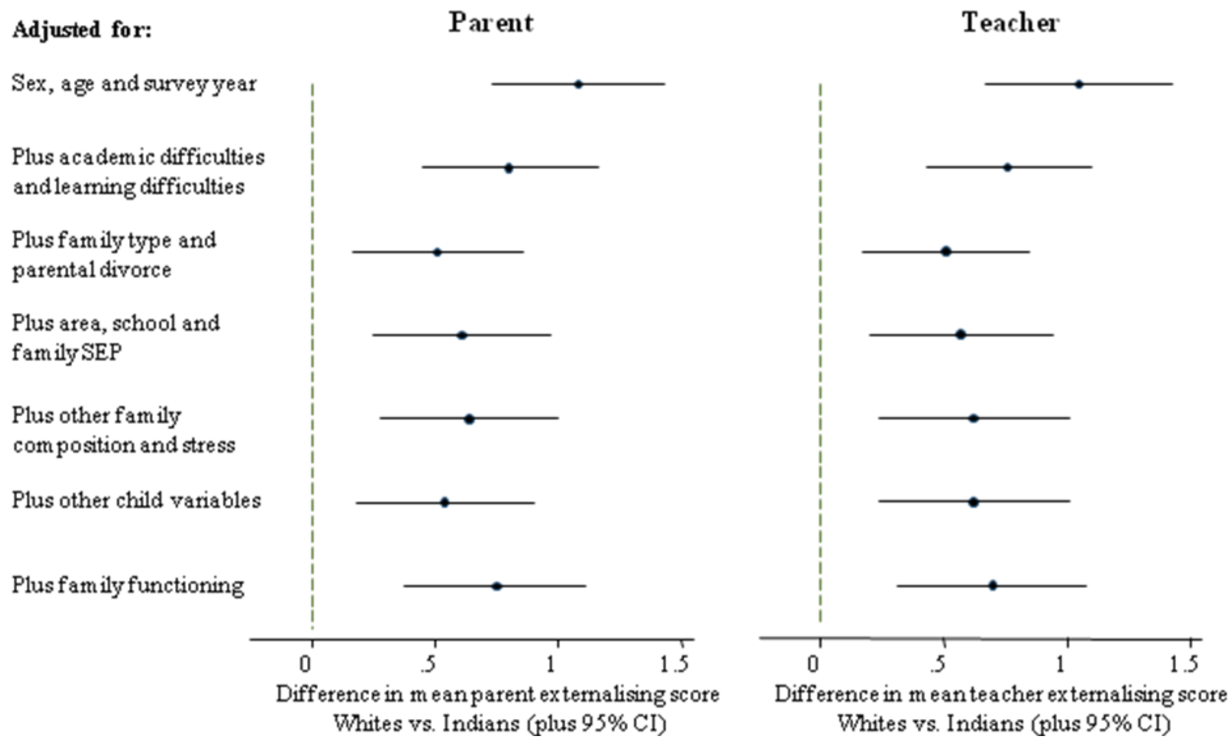
The final, fully-adjusted regression coefficients for the Indian advantage in the full population were therefore 0.75 SDQ points (95% 0.38, 1.11; $p<0.001$) for parent externalising scores and 0.70 SDQ points (95%CI 0.31, 1.08; $p<0.001$) for teacher scores. In other words, adjusting for all the child, family, school and area characteristics measured in B-CAMHS explained about a quarter of the total Indian advantage, with the difference between Indians and Whites remaining highly significant.

Table 11.14: Effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity

	Full population		Nested analysis: two-parent families	
Adjusted for:	**Parent externalising score (13 868 White, 361 Indian)**	**Teacher externalising score (10 775 White, 257 Indian)**	**Parent externalising score (9026 White, 329 Indian)**	**Teacher externalising score (7156 White, 237 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	0.62 (0.25, 0.98)**	0.60 (0.16, 1.03)**
Plus academic difficulties and learning difficulties	0.80 (0.45, 1.16)***	0.76 (0.43, 1.10)***	0.53 (0.16, 0.90)**	0.50 (0.14, 0.87)**
Plus family type and parental divorce	0.51 (0.17, 0.86)**	0.51 (0.17, 0.85)**	0.49 (0.12, 0.85)*	0.48 (0.11, 0.85)*
Plus area, school and family SEP (geographical region, metropolitan region, area deprivation, Ford score, parent education, household income, housing tenure, social class)	0.61 (0.25, 0.97)**	0.57 (0.20, 0.94)**	0.59 (0.20, 0.98)**	0.59 (0.18, 1.00)**
Plus other family composition and stress (three-generation family, no. co-resident siblings mother's age, family financial crisis, family police contact and death of parent or sibling)	0.64 (0.28, 1.00)**	0.62 (0.24, 1.01)**	0.63 (0.23, 1.02)**	0.67 (0.24, 1.09)**
Plus other child variables (neuro-developmental problems, developmental problems, common physical health problems, rare physical health problems, child hospitalisation, death of a friend, dyslexia)	0.54 (0.18, 0.90)**	0.62 (0.24, 1.01)**	0.55 (0.16, 0.94)**	0.69 (0.26, 1.11)**
Plus; family functioning	0.75 (0.38, 1.11)***†	0.70 (0.31, 1.08)***†	0.77 (0.37, 1.16)***	0.76 (0.33, 1.19)**
[Plus mother's and father's economic activity, and parent marital status: nested analysis]			0.77 (0.38, 1.17)***	0.82 (0.40, 1.25)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. †Full model presented in Table 11.17

Figure 11.10: Forest plot of the effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity (full sample)



Sensitivity analysis: further externalising and internalising outcomes

I replicated the multi-Level analyses in Table 11.14 using DAWBA diagnosis for externalising disorder and child-reported externalising score. As shown in Table 11.15, these measures closely replicated the results for the parent and teacher externalising scores. Again, adjusting for academic abilities and family type decreased the White (vs. Indian) regression coefficient somewhat but most of the difference remained unexplained. The same applied to the parent, teacher and child DAWBA bands for behavioural and hyperactivity problems (Table 14.52 Section 14.8.1, Appendix 2).

I also repeated the multi-Level analyses using DAWBA diagnosis for emotional disorder and the parent, teacher and child internalising scores (Table 11.15). As in unadjusted analyses, there was no convincing evidence that adjusting for the characteristics of Indian children unmasked an unexplained difference between Indians and Whites for internalising problems. The only instance in which the adjusted models provided any evidence of a difference was for the parent internalising SDQ score. This evidence was marginal (p=0.04) in the final model, however, and was not replicated in any of the other three internalising outcomes. It was likewise not replicated in any of the emotional DAWBA bands, including the parent DAWBA band (Table 14.52 Section 14.8.1, Appendix 2).

Table 11.15: Repeating the final multivariable analyses with the DAWBA and alternative SDQ subscales

		Regression coefficient from linear regression			Odds ratio from logistic regression
	Adjusted for:	**Parent SDQ (13 868 White, 361 Indian)**	**Teacher SDQ (10 775 White, 257 Indian)**	**Child SDQ (5737 White, 154 Indian)**	**DAWBA (13 868 White, 361 Indian)**
Externalising problems	**Sex, age and survey year**	1.08 (0.73, 1.43)***	1.05 (0.67, 1.43)***	1.24 (0.70, 1.77)***	3.98 (1.59, 9.97)**
	Plus academic difficulties and learning difficulties	0.80 (0.45, 1.16)***	0.76 (0.43, 1.10)***	1.18 (0.64, 1.72)***	3.35 (1.26, 8.94)*
	Plus family type and parental divorce	0.51 (0.17, 0.86)**	0.51 (0.17, 0.85)**	0.97 (0.45, 1.50)***	2.48 (0.93, 6.60) [p=0.07]
	Plus area, school and family SEP	0.61 (0.25, 0.97)**	0.57 (0.20, 0.94)**	0.91 (0.37, 1.45)**	2.63 (0.98, 7.10) [p=0.06]
	Plus other family composition and stress	0.64 (0.28, 1.00)**	0.62 (0.24, 1.01)**	0.91 (0.38, 1.44)**	2.59 (0.97, 6.91) [p=0.06]
	Plus other child variables	0.54 (0.18, 0.90)**	0.62 (0.24, 1.01)**	0.86 (0.34, 1.39)***	2.46 (0.92, 6.60) [p=0.07]
	Plus; family functioning	0.75 (0.38, 1.11)***	0.70 (0.31, 1.08)***	1.01 (0.50, 1.52)***	2.69 (1.00, 7.23)*
Internalising problems	**Sex, age and survey year**	-0.21 (-0.67, 0.25)	0.30 (-0.20, 0.80)	0.15 (-0.33, 0.62)	1.86 (0.89, 3.89)
	Plus academic difficulties and learning difficulties	-0.41 (-0.87, 0.04) [p=0.07]	0.11 (-0.36, 0.58)	0.08 (-0.40, 0.56)	1.67 (0.79, 3.55)
	Plus family type and parental divorce	-0.58 (-1.03, -0.13)*	0.00 (-0.47, 0.47)	-0.04 (-0.50, 0.43)	1.33 (0.62, 2.81)
	Plus area, school and family SEP	-0.49 (-0.95, -0.03)*	0.11 (-0.38, 0.59)	-0.02 (-0.49, 0.45)	1.42 (0.67, 3.03)
	Plus other family composition and stress	-0.48 (-0.94, -0.02)*	0.12 (-0.37, 0.60)	-0.01 (-0.47, 0.46)	1.42 (0.65, 3.07)
	Plus other child variables	-0.62 (-1.07, -0.16)**	0.07 (-0.41, 0.56)	-0.06 (-0.51, 0.39)	1.21 (0.57, 2.56)
	Plus; family functioning	-0.49 (-0.94, -0.03)*	0.10 (-0.38, 0.58)	-0.03 (-0.47, 0.42)	1.27 (0.60, 2.67)

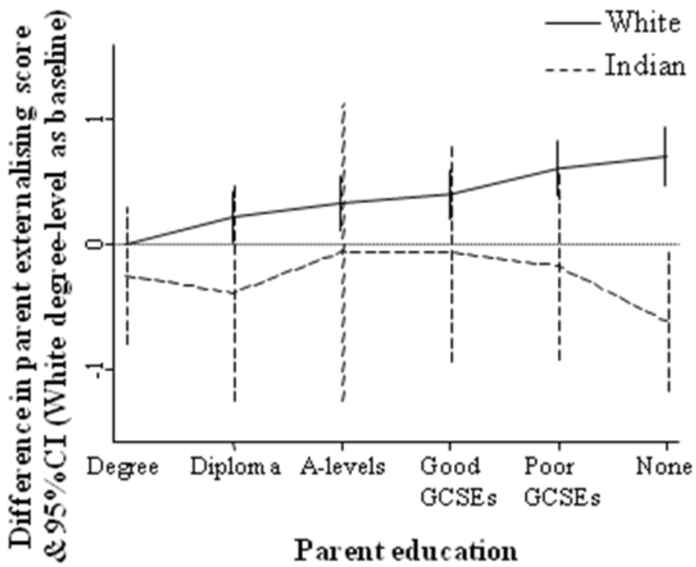
*p<0.05, **p<0.01, ***p<0.001. Table presents regression coefficients for White (vs. Indian) ethnicity from linear regression for the SDQ outcomes and logistic regression for DAWBA diagnosis.

Interaction between ethnicity and SEP in the fully-adjusted model

In Section 11.1.1, I presented evidence that all SEP/area deprivation indicators had a flatter mental health gradient in Indians than in Whites. As a result the Indian mental health advantage was largest in the more socio-economically disadvantaged groups, with little or difference from Whites in the most advantaged groups.

I used the multilevel model presented in Table 11.14 as a basis for examining how far this interaction was explained by Indians' child, family, school and area characteristics. In fact, some evidence of an interaction between ethnicity and SEP remained even after adjusting for all these variables. For example, as illustrated in Figure 11.11, in the final fully-adjusted model the significance of the interaction term between parent education and ethnicity was p=0.007 (or p=0.04 if education was entered as a categorical variable). Once again, the nature of this interaction was such that the marked SEP gradient in Whites was absent in Indians, and consequently the Indian advantage was greatest in the more deprived groups. This is also indicated by the stratified analyses in Table 11.16. As these show, the fully-adjusted regression coefficient of White (vs. Indian) ethnicity was 1.25 (95%CI 0.62, 1.88) in parents of no education, compared to 0.61 (-0.06, 1.28) in parents of GCSE-level education and 0.45 (-0.05, 0.94) in parents with A-levels or above. Moreover, this approximate three-fold difference between the bottom and the top education strata was not confined to the fully-adjusted model. Rather it was fairly constant across all the models in Table 11.16 – for example in the unadjusted model the point estimate was 2.04 for no education vs. 0. 63 for A-level education or above. This indicates that just as the measured characteristics of Indian children could not fully explain the overall Indian advantage, these characteristics also do not explain the flattening of the SEP gradient.

Figure 11.11: Regression coefficients for Indians and Whites by parent education, in predicting to the parent externalising score; based on the final fully-adjusted in Table 11.16



p-value for interaction 0.04 [0.007 if linear]

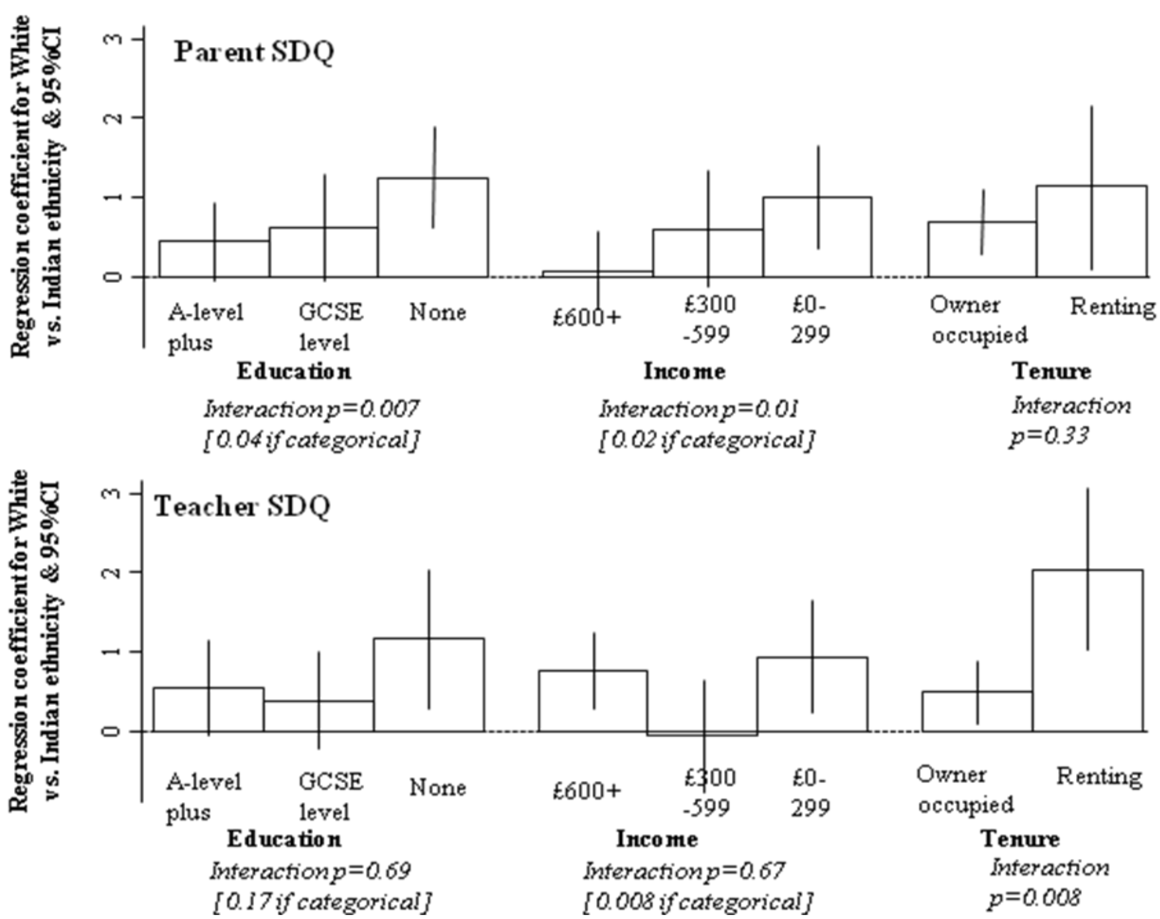
Table 11.16: Effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity; stratified analyses by parent education

	Parent externalising score				
Adjusted for:	**Full population (13 815 White, 358 Indian)**	**p-value for interaction with parent education**	**A-level qualifications or above (4698 White, 124 Indian)**	**GCSE-level qualifications (6400 White, 132 Indian)**	**No education (2717 White, 102 Indian)**
Sex, age and survey year	1.06 (0.71, 1.42)***	<0.001 [<0.001 if categorical]	0.63 (0.19, 1.06)***	0.97 (0.26, 1.69)**	2.04 (1.43, 2.64)***
Plus academic difficulties and learning difficulties	0.79 (0.43, 1.14)***	<0.001 [0.002 if categorical]	0.35 (-0.12, 0.82)	0.63 (-0.04, 1.30) [p=0.07]	1.72 (1.10, 2.33)***
Plus family type and parental divorce	0.50 (0.16, 0.84)**	0.002 [0.01 if categorical]	0.19 (-0.27, 0.66)	0.32 (-0.34, 0.99)	1.37 (0.77, 1.98)***
Plus area, school and family SEP, <u>except</u> parent education	0.57 (0.21, 0.94)**	0.008 [0.03 if categorical]	0.31 (-0.19, 0.81)	0.39 (-0.28, 1.06)	1.19 (0.53, 1.85)***
Plus other family composition and stress	0.60 (0.24, 0.96)**	0.02 [0.07 if categorical]	0.41 (-0.10, 0.93)	0.41 (-0.28, 1.09)	1.11 (0.43, 1.78)**
Plus other child variables	0.51 (0.15, 0.87)**	0.01 [0.05 if categorical]	0.31 (-0.20, 0.81)	0.34 (-0.33, 1.01)	1.00 (0.35, 1.64)**
Plus family functioning	0.73 (0.36, 1.09)***	0.007 [0.04 if categorical]	0.45 (-0.05, 0.94) [p=0.08]	0.61 (-0.06, 1.28) [p=0.08]	1.25 (0.62, 1.88)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Note that data on parent education was missing on 56 individuals, and these individuals are excluded from these analyses.

I repeated these analyses using the teacher externalising score as the outcome and using household income and tenure as SEP indicators.[37] In all cases, there was again a trend for the Indian advantage to be largest in the least advantaged group (Figure 11.12; for full models, see Section 14.8.5 Appendix 2). Likewise, the relative gap between the top and bottom groups was again similar in the fully adjusted model compared to the unadjusted model. In several cases, however, the interaction became only weakly significant or non-significant in the fully adjusted models. This was particularly true when using the teacher outcome, for which fewer individuals were available. This highlights the fact that these stratified analyses and tests for interaction are operating at the limits of the power offered by the B-CAMHS sample size. As a result, the consistent suggestion of an unexplained flattening of the SEP gradient in Indian children should be interpreted as a suggestive and hypothesis-generating trend rather than a definitive finding.

Figure 11.12: Regression coefficients from fully adjusted model for White (vs. Indian) ethnicity, stratified by parent education, household income and housing tenure



[37] These were the other two SEP indicators which showed evidence of independent predictive effects upon child mental health; see Table 11.17, p.329.

The effect of other child, family, school and area factors on externalising problems in multivariable analyses

In accordance with the aim of this PhD, the past two Chapters have focussed upon the regression coefficient for the difference between Whites and Indians. Section 14.8.3 Appendix 2 also presents in full (i.e. for all variables) the final, full-population models for the Level-specific and multi-Level models in this section. The results of the final parent and teacher multi-Level models are also presented in Table 11.17.

In addition to White ethnicity, parent externalising problems were independently predicted by male gender and younger age (but not by survey year). As in univariable analyses, geographical and metropolitan region were not predictive of externalising problems and area deprivation was not longer predictive after adjusting for family SEP. All the family SEP variables were independently predictive in multivariable analyses confined to the Level 1 variables, but only lower parental education and rented tenure remained independent predictors in the fully-adjusted model. School Ford score also remained strongly associated with externalising problems in the fully-adjusted model. Among the family composition variables, family type remained predictive of externalising problems but, unlike in univariable analyses, this was now driven exclusively by higher scores in step-families; lone parent families ceased to be disadvantaged after adjustment for family SEP. This may in part be because the relationship between lone parent families and externalising problems is confounded by low SEP but it may also reflect low SEP acting as a mediator (e.g. via post-separation declines in income). Among the other family factors, a younger mother, poorer family functioning, parental separation and family police contact also predicted externalising problems, as did all of the child variables relating to physical disorders, stressful life events and academic abilities.

The teacher model replicated the findings of the parent model except in two respects. The first was that the Ford score was highly significant in the parent model ($p<0.001$) but was non-significant in the teacher model. This is a counterintuitive finding given that one would expect school quality to have the greatest effect upon problems at school. The other difference was that the specific physical disorders and hospitalisation had much smaller point estimates of effect and were non-significant in the teacher model. These null findings seemed to result primarily from the inclusion in the models of the academic difficulties variables (teacher-reported academic abilities, parent-reported learning difficulties and parent-reported dyslexia). Removing these variables caused the Ford score to go from being highly non-significant (point estimate 0.01, p=0.62) to weakly significant (point estimate 0.32, p=0.03). Common physical problems likewise became significant (p=0.01) and developmental problems highly significant (p<0.001).

This suggests the deleterious effects of poor school quality and poor physical health may be partly mediated through academic difficulties. The teacher-reported measure of academic difficulties explains more of the variance in the teacher externalising score than parent externalising score (R-squared 23% in teachers vs. 17% in parents, in otherwise unadjusted models). By contrast, parent-reported learning difficulties and dyslexia explain less of the variance in the teacher than the parent externalising scores (R-squared 6% vs. 10%). Similarly, the estimated effect of parent-reported learning difficulties is twice as large in the parent model as in the teacher model, and vice versa for teacher-reported academic difficulties. These informant-specific effects may partly reflect shared rater biases. They may also partly genuine substantive effects – for example, children doing badly at school manifest more problems at school while children doing less well than their parents would like manifest more problems at home. In any case, the combination of informant-specific effects with the greater detail of the teacher-reported measure of academic difficulties may explain why the effect of adjustment for academic difficulties is greatest in the teacher model.

Other than these comparatively modest differences, the predictors of teacher externalising problems broadly replicated those seen in parents. This once again demonstrates the broad cross-informant consistency which has been notable throughout this Chapter.

Table 11.17: Fully adjusted models of child, family, school and area risk and protective factors for externalising problems

Domain	Variable	Categories	Parent externalising score (13 868 White, 361 Indian)	Teacher externalising score (10 775 White, 257 Indian)
Ethnicity	**Ethnicity**	Indian	0***	0***
		White	0.75 (0.38, 1.11)	0.70 (0.31, 1.08)
***A priori* confounders**	**Child's sex**	Male	0***	0***
		Female	-0.93 (-1.04, -0.82)	-1.67 (-1.79, -1.54)
	Child's age	Change per year	-0.10 (-0.12, -0.08)***	-0.03 (-0.05, -0.01)*
	Survey year	1999	0	0
		2004	-0.07 (-0.21, 0.07)	0.06 (-0.08, 0.20)
Area	**Geographical region**	South East	0	0
		London	0.00 (-0.25, 0.26)	-0.04 (-0.33, 0.25)
		South West	-0.09 (-0.33, 0.15)	0.02 (-0.22, 0.25)
		Eastern	-0.14 (-0.36, 0.08)	-0.09 (-0.32, 0.14)
		East Midlands	0.06 (-0.22, 0.34)	-0.15 (-0.40, 0.11)
		West Midlands	0.08 (-0.18, 0.33)	-0.02 (-0.27, 0.24)
		North East	0.20 (-0.18, 0.57)	-0.21 (-0.52, 0.11)
		North West & Merseyside	0.05 (-0.17, 0.28)	-0.09 (-0.33, 0.14)
		Yorkshire & Humberside	0.10 (-0.15, 0.35)	-0.13 (-0.39, 0.12)
	Metropolitan region	Non-Metropolitan	0	0
		Metropolitan	-0.01 (-0.16, 0.14)	0.12 (-0.04, 0.28)
	Area deprivation	Change per standard deviation	0.01 (-0.08, 0.10)	0.01 (-0.08, 0.10)
School	**Ford score**	Change per point	0.05 (0.03, 0.08)***	0.01 (-0.02, 0.03)
Family SEP	**Parent's highest educational qualification**	No qualifications	0***	0
		Poor GCSEs	-0.08 (-0.28, 0.13)	-0.25 (-0.52, 0.01)
		Good GCSEs	-0.27 (-0.45, -0.09)	-0.35 (-0.59, -0.11)
		A-level	-0.33 (-0.55, -0.10)	-0.27 (-0.55, 0.02)
		Diploma	-0.42 (-0.64, -0.20)	-0.25 (-0.53, 0.02)
		Degree	-0.61 (-0.84, -0.38)	-0.21 (-0.48, 0.07)
	Weekly household income	£0-99	0	0**
		£100-199	0.10 (-0.28, 0.48)	-0.21 (-0.66, 0.24)
		£200-299	0.15 (-0.24, 0.55)	-0.02 (-0.51, 0.46)
		£300-399	0.25 (-0.16, 0.67)	-0.19 (-0.68, 0.30)
		£400-499	0.30 (-0.12, 0.73)	-0.29 (-0.78, 0.21)
		£500-599	0.20 (-0.23, 0.64)	-0.26 (-0.79, 0.26)
		£600-769	0.22 (-0.20, 0.63)	-0.12 (-0.62, 0.39)
		£770 and over	0.32 (-0.09, 0.73)	0.16 (-0.34, 0.67)
	Housing tenure	Owner occupied	0***	0**
		Social sector rented	0.37 (0.18, 0.57)	0.42 (0.18, 0.65)
		Privately rented	0.28 (0.03, 0.54)	0.20 (-0.11, 0.50)
	Occupational social class	I	0	0
		II	-0.12 (-0.36, 0.13)	0.10 (-0.15, 0.34)
		III Non-manual	0.11 (-0.16, 0.37)	0.22 (-0.06, 0.50)
		III Manual	-0.08 (-0.36, 0.20)	0.03 (-0.26, 0.33)
		IV	0.03 (-0.25, 0.32)	0.13 (-0.18, 0.44)

Domain	Variable	Categories	Parent externalising score (13 868 White, 361 Indian)	Teacher externalising score (10 775 White, 257 Indian)
		V	0.01 (-0.37, 0.39)	-0.01 (-0.46, 0.44)
		Never worked	0.48 (-0.08, 1.05)	0.22 (-0.56, 1.00)
		Full-time student	-0.04 (-0.70, 0.61)	-0.02 (-0.78, 0.73)
Family composition	**Family type**	Two-parent family	0***	0*
		Step family	0.47 (0.23, 0.70)	0.33 (0.07, 0.58)
		Lone parent family	0.06 (-0.20, 0.32)	0.16 (-0.16, 0.47)
	Three generation family	No grandparent in household	0	0
		Grandparent in household	-0.16 (-0.55, 0.22)	0.22 (-0.26, 0.70)
	Number of co-resident siblings	0	0	0**
		1	0.09 (-0.07, 0.24)	-0.22 (-0.40, -0.04)
		2	0.02 (-0.15, 0.20)	-0.38 (-0.58, -0.18)
		3	0.24 (-0.02, 0.50)	0.03 (-0.29, 0.34)
		4 or more	0.00 (-0.41, 0.42)	-0.10 (-0.61, 0.40)
	Mother's age at child's birth	Change per decade	-0.05 (-0.06, -0.04)***	-0.03 (-0.05, -0.02)***
Family stress	**Family functioning**	Change per standard deviation	0.75 (0.68, 0.81)***	0.24 (0.17, 0.30)***
	Parental separation	No	0**	0***
		Yes	0.34 (0.11, 0.56)	0.12 (-0.06, 0.31)
	Family financial crisis	No	0	0
		Yes	0.15 (-0.03, 0.32)	0.12 (-0.06, 0.31)
	Family police contact	No	0***	0*
		Yes	0.52 (0.24, 0.79)	0.42 (0.09, 0.75)
	Death of parent or sibling	No	0	0*
		Yes	0.17 (-0.16, 0.50)	0.50 (0.25, 0.74)
Child	**Neuro-developmental disorder**	No	0 [p=0.05]	0
		Yes	0.62 (0.00, 1.24)	-0.19 (-0.92, 0.54)
	Developmental problems	No	0***	0
		Yes	0.94 (0.71, 1.17)	0.02 (-0.25, 0.28)
	Common physical disorder	No	0***	0
		Yes	0.39 (0.27, 0.50)	0.10 (-0.02, 0.22)
	Rare physical disorder	No	0**	0
		Yes	0.35 (0.11, 0.58)	-0.04 (-0.33, 0.24)
	Child hospitalisation	No	0*	0
		Yes	0.16 (0.00, 0.32)	0.00 (-0.18, 0.19)
	Death of friend	No	0***	0 [p=0.07]
		Yes	0.55 (0.28, 0.81)	0.26 (-0.02, 0.54)
	Teacher-reported academic difficulties	Change per point	0.34 (0.32, 0.37)***	0.61 (0.58, 0.65)***
	Learning difficulty	No	0*	0***
		Yes	1.67 (1.42, 1.92)	0.62 (0.28, 0.96)
	Dyslexia	No	0*	0*
		Yes	0.45 (0.11, 0.78)	-0.50 (-0.92, -0.08)

11.3 Discussion

These analyses indicate that family type and academic abilities play an important role in explaining the Indian mental health advantage. Adjusting for these two factors reduced the Indian advantage by about half, from 1.08 SDQ points (95%CI 0.73, 1.43) to 0.51 (0.17, 0.86) for the parent externalising score and from 1.05 (0.67, 1.43) to 0.51 (0.17, 0.85) for the teacher externalising score. Adjusting for most other child, family, school and area characteristics had little further effect upon the difference, but adjusting for the apparent Indian disadvantage in family functioning increased the unexplained difference somewhat. The unexplained Indian advantage in the fully-adjusted model was thus 0.75 (0.38, 1.11) for the parent externalising score and 0.70 (0.31, 1.08). There was, moreover, a consistent trend for this unexplained Indian advantage to be larger in more socially disadvantaged groups. For example, the estimated Indian advantage for the parent externalising score was 1.25 (0.62, 1.88) in children whose parents had no education, as compared to 0.45 (-0.05, 0.94) in children of parents with A-level education or above (p-value for interaction 0.007).

I return to these key findings in Chapter 12, where I bring together the main results from throughout this thesis and consider their implications for understanding the mental health of Indian children and of children more generally. The discussion below therefore serves a prelude to this broader treatment and focuses specifically on the results of this and the previous Chapter.

<u>Explaining the Indian mental health advantage: the role of the child, family, school and area characteristics measured in B-CAMHS</u>

Chapter 10 demonstrated that the child, family, area and school characteristics measured in B-CAMHS showed associations with externalising symptoms which were in line with the previous literature. The similarities and differences between Indians and Whites for these characteristics were also generally consistent with previous studies. Indian families showed a concentration in more deprived areas; a similar income distribution to Whites; a more bimodal distribution of education; high and relatively socially undifferentiated home-ownership; and lower mother's economic activity but proportionally higher rates of full-time vs. part-time work. This replicates the findings of other British studies, including the

UK census [295, 317], the Fourth National survey of ethnic minorities [307] and the Millennium Cohort Study [325] (reviewed in Section 3.2.2 Chapter 3). Overall, Indians were not consistently advantaged or disadvantaged socio-economically, and adjusting for SEP therefore explained very little of the Indian advantage. The Indian advantage was likewise not explained by the other area and school variables.

Yet while the Indian advantage was not explained by a more favourable socio-economic profile, there was consistent evidence that the deleterious effects of socio-economic disadvantage were less strong in Indians. As a consequence, the Indian advantage was most pronounced in more deprived groups and smaller (and often non-significant) in more advantaged groups. This flattening of the Indian SEP gradient was clearly visible in univariable analyses of all five indicators assessed: area deprivation, parent education, household income, housing tenure and social class. Some evidence for this interaction also persisted in multivariable analyses. Again the trend was always for the Indian advantage to be largest in the most disadvantaged groups – often being two to three times larger than the estimated Indian advantage in the most advantaged groups. Moreover, this two- to three-fold difference was little changed by adjusting for other child, family, school and area characteristics. The interaction between Indian ethnicity and SEP thus seemed to be largely unexplained by the variables measured in B-CAMHS.

Unfortunately, these tests for interactions and stratified analyses were operating at the limits of the power offered by the B-CAMHS sample size. My conclusions regarding the apparent interaction between SEP and ethnicity are therefore hypothesis-generating rather than definitive. Nevertheless, the consistency with which this pattern was observed across SEP indicators is striking. These findings also replicate two previous demonstrations of a flattening of the SEP gradient in Indians relative to Whites. The first was my demonstration in Section 6.3 Chapter 6 that the marked gradient in non-participation rates for White parents and children by area deprivation was not observed in Indians (see Figure 6.2, p.163). The second is a previous analyses of B-CAMHS99, which found that Whites showed a substantial SEP gradient in reading ability but that this was not observed in Indians [538].

At a minimum, therefore, these findings regarding SEP are highly suggestive. They are also of potentially great importance, as they suggest that understanding the Indian mental health advantage may shed particular light onto factors which are protective in situations of socio-economic adversity. I therefore believe that investigating whether this interaction is replicated in other samples is a research priority; to my knowledge, no previous investigation has done this.[38] If the interaction is replicated, it will be of considerable interest to investigate why Indian children may be buffered against the negative effects of low family SEP. I discuss this further in Chapter 12.

A more surprising interaction related to the weak evidence that Indians had more externalising problems in areas of higher Indian ethnic densities, while the reverse was true in Whites. This apparently deleterious effect of Indian ethnic density upon Indian mental health is inconsistent with previous studies – previous studies which in fact themselves showed highly mixed findings including linear protective effects [154], U-shaped relationships with the lowest risk in the middle [157] or inverted U-shaped relationships with the highest risk in the middle [155]. Moreover, after adjustment for the important confounder of area deprivation, there was no longer any evidence of an interaction with Indian ethnic density. There was likewise no evidence that Indian ethnic density had any main effect upon child mental health after adjustment for the other Level 1 variables. These results therefore do not support the existence of ethnic density effects in this sample, and certainly provide no evidence that ethnic density plays any role in explaining the Indian advantage.

Two-parent families were substantially more common among Indians than Whites (92.2% vs. 65.4%), as were three-generation families (14.5% vs. 1.9%). Both patterns are well-documented in the census and other large studies [295, 307, 325]. A mental health disadvantage for children in non-traditional families has been well-documented in B-CAMHS (Table 5.5 and Table 5.6, Chapter 5) and many other studies (Chapter 2, p.37).

---

[38] The only exception is RELACHS, which found no evidence of an interaction between ethnicity and eligibility for free school meals upon the child SDQ (p=0.5) 388. Stansfeld, S.A., et al., *Ethnicity, social deprivation and psychological distress in adolescents: school-based epidemiological study in east London.* Br J Psychiatry, 2004. **185**: p. 233-8.. This has limited relevance to my own analyses, however, because RELACHS simultaneously compared many ethnic groups (White UK, White Other, Bangladeshi, Pakistani, Indian, Black, Mixed or Other). It also used a crude measure of SEP which, unexpectedly, did not even show a main effect association with child mental health.

My analyses extend this by replicating the finding from the 1997 Health Survey for England [130] that low SEP seemed to explain the disadvantage for lone parent families but not for step-families. My analyses also indicated that thehigher prevalence in Indians of two-parent families partially explained the Indian advantage, reducing the unexplained difference between Indians and Whites by about a third in univariable analyses. By contrast, the higher proportion of three-generation families in Indians did not seem important. This was because, contrary to one previous study of Indians and Pakistanis [539], there was no evidence that living in three-generation families had a protective effect in either Whites or Indians.

There was no evidence of any difference between Indian and White parents for common mental health problems, which is consistent with previous studies ([243-244, 424]; see also Section 4.1.4 Chapter 4). In fact, I excluded parent mental health from my final models because my analysis of the B-CAMHS follow-up data suggested it was more an outcome than a cause of child externalising problems. This finding is intriguing given that this possibility has been comparatively neglected in the literature [148]. Reverse causal patterns of association also caused me to exclude smoking, drinking and drug use; internalising problems; the parent's use of rewards and common physical punishments except at the highest level of frequency; and relations with peers. Reverse causality is plausible for all these variables and, for substance use and participation in deviant peer groups, has some support from the literature (see Section 2.2.4). Caution is therefore needed in interpreting previous cross-sectional B-CAMHS analyses treating substance use [473] and parent reward/punishment strategies [474] as risk factors. Nevertheless, my ability to investigate causal directions was limited by the crude and/or unvalidated measures used for several putative risk factors and by the existence of only a single three-year follow-up. Further longitudinal studies applying superior measures at multiple time-points are required for more conclusive evidence regarding the relative strength of the causal effect in each direction.

Among the family and child variables which did not show reverse causality relationships with externalising problems, one unexpected finding was the strong evidence of poorer family functioning in Indian families. This finding is not straightforward to interpret. On the one hand, the family functioning GF scale's factor structure and association with parent mental health was similar between Indians and Whites. This provides some evidence that the apparent difference may reflect a real disadvantage in Indian families. Nevertheless, the GF scale has not been validated in detail in this group. Moreover while I know of no previous studies investigating British Indian family functioning, studies of other aspects of family life show no Indian disadvantage. For example, the DASH study found that Indian families spent as much or more time in family activities as Whites [413] and the Millennium Cohort Survey found that White and Indian parents were very similar in the extent to which they felt they spent 'enough time' with their infant children [540].

Furthermore, even in B-CAMHS the apparent Indian disadvantage for family functioning was not consistently replicated across other indicators of the internal quality of family life. Indian and White children were no different in their perceived social support from relatives or in the number of relatives to whom they felt close, a similarity which replicates the RELACHS study [419]. And while Indian parents rewarded their children through praise and treats less often and used frequent smacking and severe physical punishment somewhat more, they also made less frequent use of the non-physical forms of punishment. Moreover, a previous study reported no difference between Indian and White parents in their use of physical punishment and found that Indians more likely to reward good behaviour [541]. The B-CAMHS finding of poorer family functioning in Indian families should therefore be treated with caution. Certainly, however, B-CAMHS provided no evidence of a uniformly more *favourable* family environment for Indian children with respect to the variables measured, and therefore no evidence that this caused their mental health advantage.

Indeed, other than family type, only academic abilities played a large role in explaining the Indian mental health advantage. That Indians suffered fewer academic difficulties was reported by both parents (e.g. 2.9% learning difficulties vs. 8.6% in Whites) and teachers (mean difficulties score 0.13 standard deviations lower in Indians). Adjusting for these reduced the univariable difference between Indians and Whites by about 20%. This did not change substantially after additionally adjusting for the formal tests of reading and spelling in the B-CAMHS99 sample, thereby providing evidence against substantial residual confounding. The observation of fewer academic difficulties in Indians is consistent with their considerably higher attainment in school, but the causes of this high attainment are largely unknown (Chapter 3, p.85). I return to this point in the next Chapter, arguing that investigating the Indian education advantage more fully may prove central to understanding their mental health advantage.

Finally, B-CAMHS replicated the 1999 Health Survey from England [343] and other studies (reviewed in Chapter 3, p.88) in finding that Indians had poorer parent-reported general health but also fewer admissions to hospital and fewer specific disorders. There was some evidence that this discrepancy reflected a reporting bias such that Indian parents systematically made less favourable assessments of their children's general health than White parents. I therefore excluded general health from my final analyses. Insofar as the direction of the difference was for Indian children to have a physical health *dis*advantage, then in any case this could not explain the Indian mental health advantage. As regards the specific physical disorders, these did not play a major role in explaining the Indian advantage. The same was true of stressful life events.

The above discussion focuses upon identifying the most single important variables for explaining the differences between Indians and Whites. The effect of simultaneously adjusting for multiple variables is also of great interest. In the final, fully-adjusted model the regression coefficient for the Indian advantage decreased by only about a quarter and remained highly significant ($p<0.001$). This was true for both the parent and teacher models. It was also replicated using the child-reported externalising SDQ score, the parent, teacher and child DAWBA bands for behavioural and hyperactivity disorders, and DAWBA diagnosis for externalising disorder. Of equal interest is the fact that in *none* of the multivariable models in this Chapter and Chapter 10 did the Indian advantage reduce by more than a quarter to a half, and in almost none did it become non-significant at the 5% level. Strikingly, this included 'extreme case' models which adjusted only for those variables which decreased the unexplained Indian advantage. It also included models adjusting for variables like social aptitudes which may be primarily an outcome of the Indian advantage and which will therefore have tended to overestimate how much of the Indian advantage was 'explained'.

Thus this thesis presents a range of models estimating how much of the Indian advantage is explained by the variables measured in B-CAMHS. The final point estimate is about a quarter, and 'extreme case' and sensitivity analyses indicate that the true value unlikely to be more than half. The majority of the observed Indian advantage for externalising problems therefore remains unexplained by the variables collected in B-CAMHS.

By contrast, when the final, fully-adjusted model was repeated using internalising problems and disorders as the outcome there was no convincing evidence of any difference between Indians and Whites. This is of interest because some authors have hypothesised that because Indian culture places a high value upon obedience and respect, children are implicitly encouraged to channel and express their difficulties through internalising rather than externalising behaviours [421, 542]. The finding that the Indian advantage for externalising problems does *not* seem to be balanced by a corresponding disadvantage for internalising problems provides evidence against a straightforward 'redirection' model of this sort. This does not mean that Indian cultural values are unimportant: on the contrary, as discussed in the next Chapter, they may play a central role in protecting children from externalising problems. My findings do, however, suggest that any such protective effect is not part of a zero-sum game in which difficulties are merely diverted rather than prevented.

In summary, the Indian advantage is specific to externalising problems, with no apparent difference from Whites for internalising problems. The advantage exists despite poorer reported family functioning, and seems to be partly explained by a higher prevalence of two-parent families and by lower rates of academic difficulties. The latter two factors are well-documented characteristics of Indians in Britain, but more caution is needed in accepting the poorer family functioning as a genuine difference. Yet even adjusting for these and all the other established child, family, school and area risk factors collected in B-CAMHS, the greater part of the Indian advantage remains unexplained. This unexplained advantage appears, moreover, to be particularly large in children from families of low SEP. Further investigation of the Indian mental health advantage therefore has the potential to identify factors which not only improve the mental health of children in general, but which also have a particularly large effect on children at risk.

Chapter 12 Final discussion and conclusion

12.1 Summary of PhD findings

The British Child and Adolescent Mental Health Surveys (B-CAMHS) of 1999 and 2004 found a substantially lower prevalence of any child mental disorder in Indians compared to the general population (3.4% vs. 9.4%, $p<0.001$). This PhD sought to understand this apparent Indian mental health advantage through secondary analyses comparing the 16 449 White and 419 Indian children aged 5-16 in B-CAMHS. The analyses in this thesis revealed strong evidence ($p\leq 0.002$) of an Indian advantage for externalising problems/disorders on all mental health measures, and little or no difference for internalising problems. This was consistently observed for clinical diagnosis and also for questionnaire measures (the SDQ) administered separately to parents, teachers and children. It was also consistent with the findings of my systematic review. Detailed psychometric analyses of the questionnaire and clinical interview measures provided no evidence that measurement bias could account for this observed Indian advantage. There was likewise no evidence that the Indian mental health advantage could be explained by participation bias.

The Indian advantage is therefore specific to externalising problems and is unlikely to be explained by either chance or bias. Part of the explanation for this differences for externalising problems seemed to be the fact that Indian children were more likely to live in two-parent families (92.2% vs. 65.4%) and less likely to have academic difficulties (e.g. 2.9% vs. 8.6% for parent-reported learning difficulties). Adjusting for these two factors in multivariable analyses reduced the unexplained difference between Indians and Whites by about half; from 1.08 to 0.51 SDQ points for the parent SDQ and from 1.05 to 0.51 points for the teacher SDQ. Most other child, family, school and area characteristics had little further effect with the exception of family functioning. Indians had poorer family functioning scores (by 0.27 standard deviations) and adjusting for it therefore increased somewhat the unexplained difference between Indians and Whites. The unexplained Indian advantage in the final fully-adjusted model was thus 0.75 (95%CI 0.38, 1.11) SDQ points for the parent externalising score and 0.70 SDQ points (95%CI 0.31, 1.08) for the teacher externalising score.

In both unadjusted and adjusted models, the unexplained Indian advantage for externalising problems was consistently larger in families of low socio-economic position (SEP). For example, in the fully-adjusted models predicting to parent externalising score, the Indian advantage was 1.27 points in children whose parents had no educational qualifications vs. 0.45 in children whose parents had A-level qualifications or above (p-value for interaction 0.007). As in unadjusted analyses, there was little or no evidence of an ethnic difference for internalising problems/disorders in adjusted models.

To summarise, the Indian mental health advantage seems to be genuine, and is specific to a reduced prevalence of externalising problems/disorders. A higher prevalence of two-parent families and academic abilities seem to mediate part of this advantage, but most of the difference is not explained. Likewise unexplained is the fact that the Indian advantage appears to be particularly large in socio-economically disadvantaged groups.

12.2 PhD strengths and limitations

Before discussing the implications of these findings, it is worth reviewing the strengths and limitations of this PhD. This PhD presents the first in-depth analysis of the validity of any widely-used child mental health measure across two ethnic groups in Britain. It also represents one of the first studies to examine in depth the causes of observed ethnic differences.

This comprehensive analysis was made possible by several central strengths of B-CAMHS. The B-CAMHS surveys are the largest and most comprehensive surveys of child mental health ever conducted in Britain, containing a nationally-representative, population-based sample of 18 415 children. Data was collected from parents, teachers and children and, uniquely among surveys of this size, was used to generate multi-informant clinician-rated diagnoses for all children. The availability of these 'gold standard' outcomes gives considerably more weight to the conclusion that Indians have a mental health advantage than was possible in previous population-based studies, all of which used only questionnaires. The SDQ questionnaires collected in B-CAMHS were nonetheless also highly important in allowing me to demonstrate that the Indian advantage for externalising problems was observed across informants, including the external informant of teachers. The SDQs were also central to my demonstration that internalising and externalising problems were meaningful constructs in both Indians and Whites. I investigated this issue of cross-cultural validity in more detail than is typical, and this gives further confidence to my conclusion that the Indian advantage is genuine. Finally, the B-CAMHS surveys collected an unusually wide range of potential explanatory variables. This allowed examination of multiple possible explanations for the apparent Indian advantage, thereby making an important advance on most previous studies which simply described ethnic differences. In investigating the causes of the Indian advantage, the B-CAMHS follow-up component was also valuable in evaluating which variables showed evidence of reverse causality

.

Nevertheless, both the B-CAMHS data and the analyses presented in this PhD have important limitations. Although taking a large sample overall, B-CAMHS did not oversample minority ethnic groups. Consequently even Indians, Britain's single largest minority ethnic group, have only a moderate sample size (N=419). For most analyses in this PhD this sample size proved adequate. Inadequate power did, however, prevent the application of some techniques which I had found useful in investigating potential reporting biase in the conceptually similar problem of the apparent mental health advantage of Norwegian vs. British children ([230], Appendix 3). Interpretation of some of my substantive analyses in Chapter 10 and Chapter 11 was likewise complicated by lack of power, with this applying particularly to the interaction between Indian ethnicity and SEP.

B-CAMHS did not oversample minority ethnic groups because investigating ethnic differences was not one of its primary purposes. This also explains the absence in B-CAMHS of some child and family information which might have been particularly informative in this PhD. For example, no information was gathered on alternative aspects of ethnic identity (e.g. religion) or within-Indian heterogeneity (e.g. East African vs. non-East African migration to Britain, or second vs. third generation children). Such information might have helped direct the focus of further investigations and generated hypotheses about the mechanisms of effects (although in the absence of oversampling, such analyses would most likely have been underpowered). B-CAMHS likewise collected no information on child and/or parent acculturation, which other studies have found to be informative when examining ethnic differences in child mental health [387, 421, 430, 543-544]. More generally, B-CAMHS provided relatively little information about the quality of the child's interaction with their family members. For example, unlike the DASH study [413, 545] or the Millennium Cohort Survey [546], there were no questions on how much time the child spent in different sorts of family activities. In fact, what family environment information was collected in B-CAMHS did not seem important in explaining the Indian advantage. Nevertheless, a greater involvement in family activates among Indian children did seem to explain a modest part of the advantage of Indians over Whites in DASH [413],[39] and it is therefore unfortunate that this information was not available for my analyses.

Finally, my focus in this PhD is relatively narrow, concentrating on comparing Indians and Whites for common child mental health problems. The focus on Indians and Whites is partly pragmatic in that these are the two largest ethnic groups in Britain and therefore the best powered for comparison. The focus is also justified on theoretical grounds by the fact that the Indian mental health advantage over Whites is the ethnic difference for which there is most evidence in the previous literature. It does, however, mean that my PhD deals with only one aspect of the relationship between ethnicity and child mental health in Britain. Similarly, my focus on the common child mental problems is partly pragmatic in that these *are* more common problems – and of course that is itself also a justification for considering them of particular public health importance. Moreover, it is for these that the previous literature provides the strongest evidence of an Indian advantage, with little or no evidence of differences for rarer outcomes like psychosis or deliberate self-harm ([364], Appendix 3).

[39] The Millennium Cohort Survey has not yet been used to investigate mental health.

12.3 Interpretation and implications of PhD findings

This PhD has demonstrated that the Indian advantage is unlikely to be explained by chance or bias, but rather seems to represent a genuine advantage for externalising problems/disorders. One immediate implication for child mental health practitioners is that an 'under-representation' of Indian children at mental health clinics does not necessarily reflect unmet need. It may instead reflect a lower prevalence of problems.

This PhD has also sought to investigate the causes of this Indian advantage, with a view to learning lessons which may help improve the health of all. Before discussing further my findings on this issue, it is worth reviewing why greater understanding of protective factors against externalising problems and disorders is of substantial public health interest. Externalising (particularly behavioural) problems have been increasing in Britain in the past 30 years [6-7], such that 6% of children in B-CAMHS received a diagnosis of an externalising disorder. Externalising disorders are, by definition, associated with substantial distress and impairment to the child at the time. Long-term follow-up of British birth cohorts also indicates associations with adverse future life experiences across a wide range of domains, including work life, socio-economic position, inter-personal relationships, and health [7]. Child mental health is now rightly considered a government priority in Britain, and the past decade has seen a range of new prevention initiatives [8].

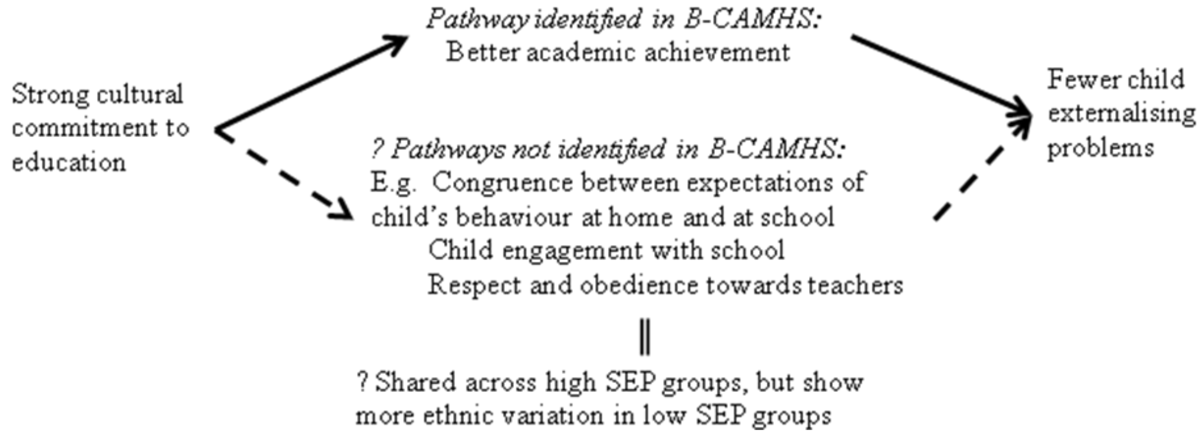
Many of the leading prevention initiatives, including SureStart and the Healthy Schools program, aim to foster good child mental health by enriching children's educational experience. In this context it is intriguing that better academic abilities do indeed seem to be important in protecting Indians from mental health problems and explaining part of their advantage over Whites. Understanding the Indian education advantage could therefore clarify a mechanism for promoting child mental health which is of great political interest.

What, then, is known about the Indian education advantage? As reviewed in Chapter 3 (p.85), this advantage is certainly well-documented. Indian children consistently out-perform White British children in Key Stage tests at every age [330] and also make better progress between Key Stages [332]. At GCSE level the proportion of Indians getting five good GSCEs is almost 15% higher than in White British students (61.6% vs. 47.3% for boys and 71.9% vs. 57.3% for girls) [331] and Indian 16-24 year olds are almost twice as likely to hold degrees as Whites (19.6% vs. 10.4%) [326]. Yet surprisingly and unfortunately, the cause of this high attainment in Indians has received very little attention in educational research. Some qualitative surveys highlight the importance of education among Indian families [333, 547], but interpretation of these is complicated by the lack of comparison groups. Recent quantitative surveys have focussed on disadvantaged minority groups [336] and/or using meta-ethnic categories like 'Asian' [337].

Yet while more research is clearly needed, the existing literature does contain some findings regarding 'South Asian' samples which are at least suggestive. One which I find particularly intriguing comes from a qualitative survey of London parents which suggests alternative routes whereby ethnic differences in parental attitudes towards education may affect child mental health. In this study, South Asian (mostly Indian and Pakistani) parents stressed success in school as a route to greater confidence and self-advancement, while White parents more often said they wanted their child to do as well as they could and to enjoy school [214].

One research priority is therefore further investigation of why Indians do so well in the education system. Such research would undoubtedly be valuable in clarifying the origin of one *identified* mechanism for the Indian mental health advantage, namely their greater academic abilities. It might also generate insights or hypotheses regarding as yet unidentified mechanisms, and so shed light on the portion of the Indian mental health advantage which is currently unexplained. As shown schematically in Figure 12.1, I speculate that a strong cultural commitment to education may lead to attitudes and practices which have additional protective effects against externalising problems. These could, for example, include a greater congruence between how the child is expected to behave at school and at home, or giving the child a sense that being at school is meaningful rather than a waste of their time. A strong commitment to education may also intersect with, and be reinforced by, a more general cultural emphasis upon respect and obedience towards adult authority figures such as teachers. I return to this point in the next Section.

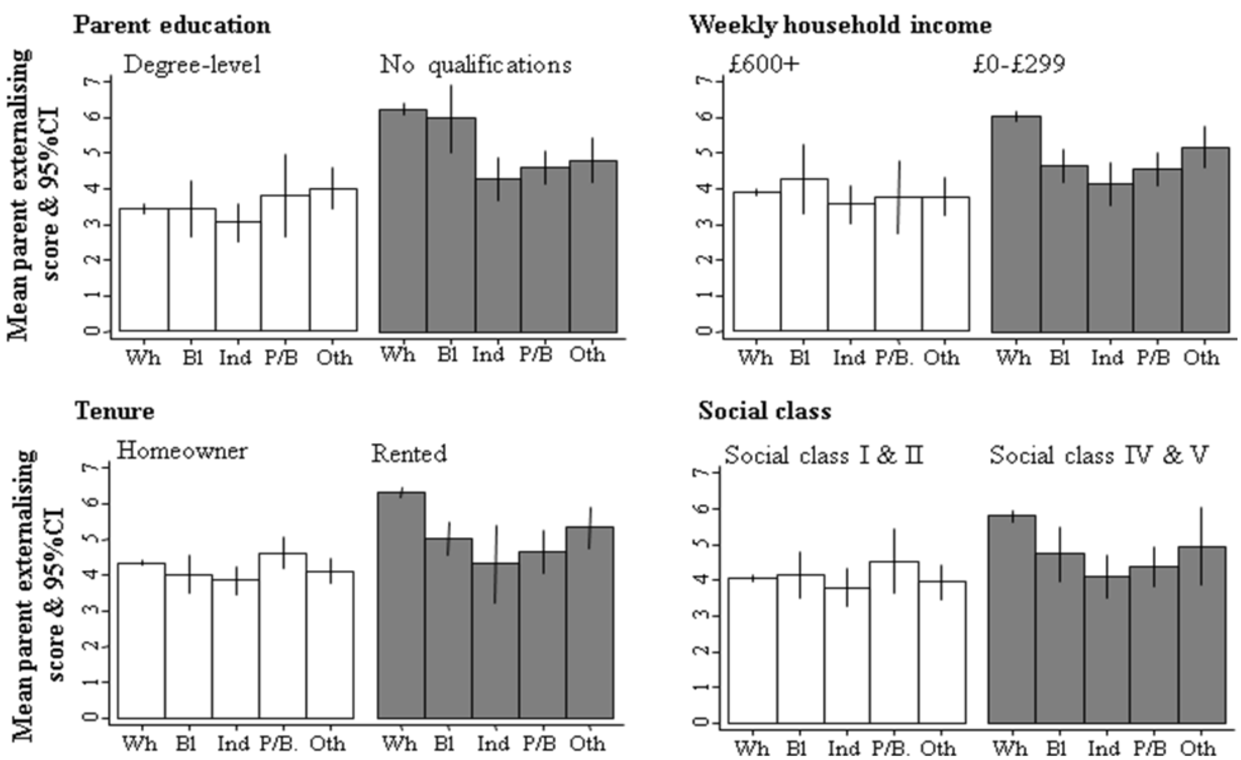
Figure 12.1: Hypothesised additional mechanisms whereby a cultural commitment to education could protect against child externalising problems



Understanding these additional pathways might also generate insight into why the Indian mental health advantage seems to be particularly large in families of low SEP, with little or no ethnic difference in families of high SEP. If replicated in other datasets, this interaction with SEP will be of great interest and importance. The fact that there is little or no Indian advantage in high SEP groups provides some evidence that the Indian advantage may not stem from something which is confined to Indians, such as protective gene alleles or a highly culturally-specific set of values. Instead the advantage may reflect a set of attitudes and behaviours which have the potential to exist across ethnic groups, but which currently in Whites are largely confined to high SEP families. This may, moreover, be closely related to the unmeasured protective pathways hypothesised in Figure 12.1. Certainly attitudes towards education are a plausible candidate for explaining part of the marked SEP gradient in Whites, given the long-standing education-orientation of the White middle-classes and the ambivalence and resistance which these values evoke in some working class populations [548]. Moreover, there are at least some suggestions that such class differences in attitudes to education may be less marked in Indians. For example, an early qualitative study in British Sikhs emphasised that high education aspirations were consistently observed across different social and occupational classes [547]. A more recent study likewise reports that there has been an "adaptation of middle-class values towards education by working-class South Asians" [549, p.304]. This also resonates with Modood's characterisation of the Indian integration strategy as 'waiting to assimilate' until achieving entry to the middle classes ([241]; see Chapter 3, p.79). Thus even Indian families of low income or education may still show a ‘middle-class’ orientation to education which protects their children’s mental health.

I further hypothesise that this interaction with SEP may extend to other ethnic groups in Britain. Specifically, I hypothesise that socio-economically advantaged families share constellations of protective attitudes or practices which transcend ethnic boundaries, whereas attitudes and practices in disadvantaged families show far more ethnic variation. If so, this would echo the observation that all-cause mortality in Britain shows massive geographical variation in social class V but little or no difference geographical variation in social class I [550]; or that literacy scores show substantial between-country variation among young people whose parents have low education but tend to converge among those whose parents have high education [551]. Exploratory analyses in B-CAMHS provide preliminary support for my hypothesis, showing relatively small ethnic differences in parent-reported externalising scores among high SEP families as contrasted with substantially larger ethnic differences between low SEP families (Figure 12.2). All these SEP-ethnicity interactions were highly significant (p<0.001).

Figure 12.2: Ethnic differences in parent externalising scores among low SEP and high SEP families



Wh=White, Bl=Black, Ind=Indian, P/B=Pakistani and Bangladeshi, Oth=Other ethnicity. Based upon all children from England, Scotland and Wales with non-translated parent SDQs (N=18 223)

These findings regarding other minority ethnic groups are of course preliminary and my suggestions regarding the possible role of attitudes towards education are speculative. Yet whatever the mechanism, and even if the interaction does not generalise across ethnic groups, the apparent flattening of the SEP gradient in Indians is still of substantial importance. It indicates that low SEP does not inevitably reduce child mental health, and suggests that the unknown factors which explain the Indian advantage may generate insights which are particularly relevant for reducing the marked SEP gradient in White children.

Using these insights to reduce this gradient would, of course, be a formidable challenge. Nevertheless, socio-economic gradients in positive parenting behaviours are not immutable – for example, recent findings from the YouthTrends survey show that while in 1986 there was a strong social class gradient in parental monitoring and supervision, this had disappeared by 2006 [552]. As such, not only may understanding the Indian advantage hold important lessons for improving child mental health in general, it may also hold some clues for promoting child mental health equity in particular.

12.4 Further directions for future research

This PhD has made some important contributions to what we know about the mental health of Indian children. Most of the Indian advantage for externalising problems was, however, not explained by the many known risk factors collected in the B-CAMHS surveys. My systematic review likewise found that although other studies consistently observed the Indian advantage, they had largely failed to explain it.

This PhD therefore does not fully ‘solve’ the mystery which motivated it. I believe that this failure is itself a finding of substantial interest. Taken in combination with the compelling evidence that the Indian advantage is real, the fact that the advantage is not wholly explained by standard risk factors confirms that further investigation may generate genuinely new insights about protective factors. In the previous section, I argued that further investigation of the Indian education advantage might prove central to achieving this goal. There are also other valuable directions for further research, including both qualitative and quantitative studies, and including both exploratory and hypothesis-driven investigations. In this final Section, I outline what I believe should be the immediate priorities.

Qualitative research into explanatory models of child behaviour

'Explanatory models' refer to what individuals see as the nature and causes of different states, and how they think they should react [553]. I believe that the explanatory models of Indian parents regarding children's behaviour may be important to understanding the Indian mental health advantage.

One core component of these explanatory models may be an emphasis upon obedience, respect towards adults and good behaviour. These have been highlighted as key aspects of British Indian child rearing in several quantitative, qualitative and ethnographic studies [542, 554-555], although unfortunately few studies have included an explicit mental health focus. The one notable exception is the comparison by Hackett *et al.* of 100 Gujarati Indian parents and 100 Whites [10, 541, 556]. This reported that the Indian parents showed higher expectations of obedience and self-control, being less tolerant of disruptive play or temper tantrums. The study also provided some hints of the potential explanatory models underlying these differences. For example, parents were asked what they would think if their child retaliated when another child behaved badly. Many Whites described this as 'self-assertive' while Indian parents were more likely to see it as 'worryingly aggressive' and preferred their child to tell a teacher.

Further suggestive evidence regarding the explanatory models underlying different parenting practices comes from the qualitative study by Nikapota *et al.* of 60 White British and 60 'South Asian' parents, together with 60 Black-Caribbean and 40 Mixed race [214]. The use of the meta-ethnic group 'South Asian' complicates interpretation of the findings, but it is nonetheless suggestive that South Asian parents were found to be more likely to stress the need for *unconditional* obedience in their children. By contrast, parents of White and Mixed race children focussed more often on the need for self-determination, while Black-Caribbeans typically qualified that obedience was only desirable for 'reasonable demands'.

I therefore hypothesise that cultural norms and parenting practices which emphasise unconditional obedience to adult authority figures may be an important cause of the Indian advantage. Moreover, this cultural norm may intersect with the factors which seem to be most important in explaining the Indian advantage. Given the importance of family type, it is notable that Hackett *et al.* state that “the relatively secure families in which these [Indian] children grow up enable their stringent expectations of behaviour to be implemented effectively” [10, p.103]. Likewise, the role of academic abilities makes it particularly salient that Nikapota *et al.* found the clearest inter-ethnic differences in expectations of obedience concerned obedience at *school*. As I suggested in the previous section, the combination of a strong cultural commitment to education and an emphasis upon obedience to adults may be central to the Indian mental health advantage.

In addition to these hypothesised differences in expectations, I also believe that there may be important differences between how Indian and White parents *respond* to challenging child behaviours. One intriguing possibility is that Indian parents may combine a low tolerance of disruptive behaviours with a ‘distress’ model for those behaviours which do occur. Thus Hackett *et al.* report that while White parents saw tantrums as essentially confrontational, Indians emphasised likely distress of the child [10]. This may help explain why the Indian parents were also more likely to respond to misbehaviour by ‘explaining’ to the child rather than shaming or teasing them [541]. This finding was not, however, replicated by Nikapota *et al*. Rather South Asians were less likely than other groups to discuss the child’s emotional state when given a vignette of a nine-year old girl who was fighting. This non-replication, in combination with the difficulty of interpreting findings from the ‘South Asian’ sample, highlights the need for further investigation.

Further investigation is also necessary because the existing qualitative data is limited to brief interviews and simple thematic analyses. To be most informative, such research should have an explicit focus upon explanatory models of child mental health. This could involve investigating respondents’ perceptions of what constitutes both good and bad mental health; what causes these states; and how one can best promote good mental health and reduce problems. Such research would ideally involve methods such as in-depth interviewing or ethnographic studies. It would also ideally go beyond descriptive accounts of multiple specific differences and instead seek to unite these within a broader analytic framework. Finally, given the intriguing but unexplained interaction between Indian ethnicity and SEP, future qualitative studies might wish to include comparisons across socio-economic groups as one specific axis of their research analysis. This might involve oversampling individuals from the extremes of the SEP distribution, in order to generate the most informative contrasts possible.

Replication and extension in quantitative surveys

There exist several studies of child mental health which contain large numbers of Indians and which could build upon the analyses in this PhD. Of particular interest is the Millennium Cohort Study, a nationally-representative cohort of infants born in 2000-2001. This oversampled for minority ethnic groups such that the English sample in the first sweep contained 8664 White mothers and 458 Indian mothers [557]. Parent SDQs were collected in the second (2004/5) and third (2006) sweeps, and the fourth sweep (2008) additionally collected teacher SDQs [558]. In addition, a wide range of other information on family life has been collected, including information on parenting styles and practices, on how much contact the child has with other family members, and on how often the child participates in various family activities [546]. No analyses have yet been published on the SDQ data, but it will clearly be of considerable interest to investigate whether the Indian advantage for externalising problems is again observed and, if so, whether these additional variables seem important in explaining the difference.

It would also be of substantial interest to investigate whether the Millennium Cohort Survey replicates the interaction between Indian ethnicity and SEP reported in this PhD; other existing studies such as DASH [413] and the Health Surveys for England [343, 393] could also be used for this purpose. If this interaction is replicated then, as discussed above, this could have important implications for child mental health equity. Specifically, understanding why Indians do *not* show a marked SEP gradient might illuminate why White children do show a strong SEP gradient and suggest how that gradient could be reduced.

In addition to making the best possible use of existing data, there is also a role for novel quantitative surveys. These should address the limitations of B-CAMHS by oversampling all minority ethnic groups; collecting information on additional axes of ethnic identity such as regional origin, religion or route of migration; and collecting information on issues such as cultural identity or the experience of racism. Given the apparent importance of academic abilities in the Indian advantage, these points also apply to future educational surveys. In particular, I believe recent surveys by the (then) Department for Education and Skills were not justified in using the above-average academic attainment of Indian and Chinese children as a reason for excluding them from the minority ethnic oversampling [336]. As I have sought to demonstrate in this PhD, studying groups with an advantage may be as informative as studying those with a disadvantage.

Ideally, future surveys will be preceded by qualitative research and can therefore test the hypotheses generated in the qualitative analyses. A follow-up component would also be valuable in allowing some assessment of how far the Indian advantage reflects a lower incidence of problems and how far it reflects faster recovery (the Millennium Cohort Study will also allow some scope to examine this).[40] The Department of Health is currently considering commissioning a new B-CAMHS survey as a longitudinal internet panel, possibly with over-sampling of minority ethnic groups. If this goes ahead, it will provide an excellent opportunity to apply these recommendations. Another excellent opportunity will be the proposed 2012 UK birth cohort which, if funded, is likely to include ethnicity as one of its central themes.

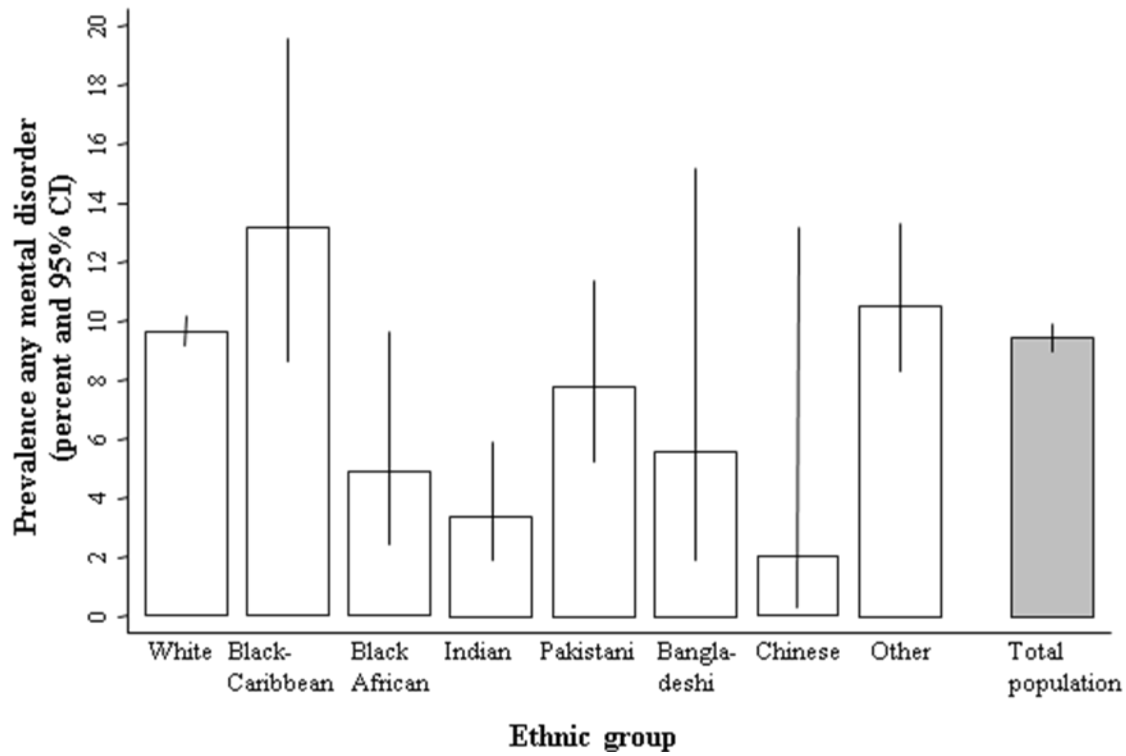
Extension to other minority ethnic groups

Finally, a study following the above recommendations would provide a platform for extending the methodologies developed in this PhD to the study of other minority ethnic groups. As outlined above, this PhD focuses upon Indians partly because they are the only group in B-CAMHS with sufficient power for detailed inter-ethnic comparisons. Yet my systematic review found a few studies which also suggested a possible mental health advantage for Black African children. B-CAMHS provides marginal evidence of such an advantage, with an odds ratio for 'any mental disorder' in Black Africans (vs. Whites) of

[40] As discussed in Section 11.1.2 Chapter 11, more frequent follow-up could help clarify the direction of causality between child mental health and other child or family variables.

0.48 (95%CI 0.23, 1.00; p=0.05; see Figure 12.3). Given that Chinese children are the other British ethnic group with a well-documented advantage in educational attainment, it is particularly intriguing that they too show a trend towards a large advantage (OR 0.20; 95%CI 0.03, 1.43; p=0.11). It is likewise intriguing that, as presented previously in Figure 12.2, the interaction between SEP and ethnicity may extend to other ethnic groups beyond Indians and Whites.

Figure 12.3: Prevalence of any mental disorder by ethnicity in B-CAMHS (eight-way classification of ethnicity)



Conclusion

One priority for future research is therefore to use oversampling to include adequate sample sizes of smaller ethnic groups, and to allow sufficient power for investigation of interactions as well as main effects. Ideally such research would also proceed in consultation with adults and young people from community organisations from the minority groups in question, in order to harness their insights and ensure the research addressed any issues of particular concern. The methodological and conceptual tools developed in this PhD could then be applied to multi-ethnic comparisons, and this would offer substantially greater scope for testing hypotheses and making informative contrasts. In combination with relevant qualitative research, this would represent a research program which rose to the challenges of making cross-cultural comparisons in mental health. These comparisons might then provide a springboard for identifying new ways to promote mental health and mental health equity in children of all ethnicities.

Why do British Indian children have an apparent mental health advantage?

Anna Goodman

London School of Hygiene and Tropical Medicine
University of London

Thesis submitted for the degree of PhD, September 2009

Part 2: Appendices and references

Chapter 13 Appendix 1: Statistical concepts and methods

Index of statistical concepts	
Clustered sampling	Section 13.3.1
Communality	Section 13.1.5 (Multi-group CFA)
Comparative Fit Index (CFI)	Section 13.1.5 (Confirmatory factor analysis)
Complex survey design	Section 13.3
Confirmatory factor analysis (CFA)	Section 13.1.5
Construct validity	Section 13.1.1
Content validity	Section 13.1.1
Convergent validity	Section 13.1.1
Cook's distance.	Section 13.2.1 (Assumptions of linear regression)
Criterion validity	Section 13.1.1
Cronbach alpha	Section 13.1.3
DEFF	Section 13.3.1 (Clustered sampling)
DEFT	Section 13.3.1 (Clustered sampling)
Design effects from clustering	Section 13.3.1 (Clustered sampling)
Discriminant validity	Section 13.1.1
Eigenvalue	Section 13.1.5 (Principal component analysis)
Exploratory factor analysis (EFA)	Section 13.1.5
Factor analysis	Section 13.1.1 and Section 13.1.5
Factor loadings	Section 13.1.5 (Exploratory factor analysis)
Finite population corrections	Section 13.3.2
Fisher's z-transformation	Section 13.1.2
Forced entry model selection	Section 13.2.3
Generalised linear models	Section 13.2.2
Geomin rotation	Section 13.1.5 (Exploratory factor analysis)
Group differentiation	Section 13.1.1
Heterotrait blocks	Section 13.1.4 (Multitrait-multimethod analyses)
Heterotrait, monomethod triangles	Section 13.1.4 (Multitrait-multimethod analyses)
Hypothesis testing	Section 13.1.1
Influence	Section 13.2.1 (Assumptions of linear regression)
Internal consistency	Section 13.1.1 and Section 13.1.3
Inter-rater reliability	Section 13.1.1
Intraclass Correlations	Section 13.1.2
Kappa statistic	Section 13.1.2
Kendall tau	Section 13.1.2
Latent variable/factor	Section 13.1.5
Leverage	Section 13.2.1 (Assumptions of linear regression)
Likelihood ratio test statistic	Section 13.2.3 and Section 13.3.2
Linear regression	Section 13.2.1
Logistic regression	Section 13.2.2
Manifest variable	Section 13.1.5
Measurement invariance	Section 13.1.5 (Multi-group CFA)
Missing At Random (MAR)	Section 13.4.1
Missing Completely At Random (MCAR)	Section 13.4.1
Missingness mechanism	Section 13.4.1
Multi-group confirmatory factor analyses	Section 13.1.5
Multinomial logistic regression	Section 13.2.2
Multiple imputation	Section 13.4
Multiple imputation by chained equations (MICE)	Section 13.4.2
Multitrait-multimethod (MTMM) analyses	Section 13.1.4
Not Missing At Random (NMAR)	Section 13.4.1
Oblique geomin rotation – *see Geomin rotation*	

Index of statistical concepts

13.1 Reliability and validity

13.1.1 Psychometric concepts of reliability and validity

Initially developed in the fields of education and psychology, psychometrics encompasses a broad set of theories and methods about the measurement of human characteristics. Reliability and validity are two core concepts in the psychometric evaluation of any measure. Reliability is defined with reference to the assumption that observed ('manifest') test scores will depend both upon a true ('latent') score and also upon random measurement error. Reliability is higher when the error variance makes up a smaller fraction of the test score variance [559]. In principle higher reliability would ideally be assessed by the correlation of independent repetitions of the same test under the same circumstances. In practice this is not possible, meaning that instead one has to use approximations based upon similar administrations of the same test and/or using data available from the covariance matrix of items in a single test administration.

Validity refers to the extent to which measures provide an unbiased and accurate measure of the construct which they claim to capture [495]. Reliability is a necessary precondition for validity as if random error is a large component of a measure then it cannot be providing a valid measure of any construct. Reliability is not sufficient for validity, however; it is possible reliably to measure the wrong thing. In other words, it is possible to make a highly replicable measurement of an underlying construct which is different to the one which you intend to measure (e.g. a claimed measure of intelligence is in fact measuring concentration levels or general knowledge).

The majors form of reliability and validity used in psychometrics are summarised in Table 13.1 and Table 13.2. These Tables describe each form of reliability and validity, and summarise its applicability to the psychometric evaluation of brief questionnaires and diagnostic interviews. These also list some of the statistical techniques used in their evaluation and which are described in full detail in Section 13.1.2.

Table 13.1: Types of reliability assessed in psychometric analysis

		Applicable to brief questionnaires?	Applicable to diagnostic interviews?	Example of statistical methods used
Test-retest reliability	Agreement between the results of a measure independently administered to the same person two or more times (the ‘test’ and the ‘retest(s)’). For most psychometric tests it is not possible to administer the retest immediately because people would remember their responses from the first time and so the retest would not be independent. As such, a gap of around two weeks is often used. For traits such as mental health which may show genuine fluctuation even over this short period, this gap is therefore expected to generate conservative estimates of test-retest reliability. This is because some of the some disagreement between the test and retest may be due to genuine changes.	Yes	Yes	ICC, Pearson’s or, Spearman’s correlation
Inter-rater reliability	Agreement between the results of a measure independently administered to two or more raters reporting on the same subject (e.g. two clinicians diagnosing the same child).	Yes	Yes	ICC, Pearson’s or, Spearman’s correlation
Parallel forms reliability	Agreement between the results of alternate forms of the same measure. This is particularly relevant for tests of knowledge or intelligence (as opposed to opinions or attitudes), where one cannot use identical questions more than once as people learn the answers. Although there are alternate forms for some measures of mental health and personality, the rationale for parallel forms is generally less relevant in these areas. Parallel forms do not exist for any of the measures used in this thesis.	Not usually in mental health	Not usually in mental health	
Internal consistency	Degree of interrelatedness of a set of items in the same scale. This is usually assessed with reference to the desire that different items on the same scale measure the same thing, but not have such a strong intercorrelation that there is item redundancy. The concept of internal consistency is therefore only readily applicable to measures which use multiple items to measure a single construct, and not to binary decisions such as whether to give a child a diagnosis for a mental disorder.	Yes	Not for diagnosis, but may be assessed for questions within a subsection	Cronbach alpha (but see Section 13.1.3 for problematic aspects)

Adapted from Lamping *et al.* [560-561], Hilari *et al.* [562] and Smith *et al.* [563]

Table 13.2: Types of validity assessed in psychometric analysis

		Applicable to brief questionnaires?	Applicable to diagnostic interviews?	Example of statistical methods used
Content validity	Whether a measure covers all the important domains or subdomains of a phenomenon of interest. This links to the conceptual framework underlying the development of a measure, and may be informed by qualitative work. It is also connected to what is called 'face validity', which is the principle that it should seem plausible to an informed reader that the measure does capture the phenomenon of interest.	Yes	Yes	Not quantitative
Criterion validity	How well a measure performs when compared against a gold standard measure ('criterion'). As argued in Section 2.4, Chapter 2, true gold standard measures do not exist in child mental health. One may, however, use the methods of assessing construct validity to compare one measure to another existing measure of known validity (e.g. a brief questionnaire vs. a detailed diagnostic interview).	No true gold standards, so usually evaluate construct validity	No true gold standards, so usually evaluate construct validity	ICC, Pearson's or Spearman's correlation, Sensitivity, specificity, ROC
Construct validity	How well a measure performs when compared to other validated measures but non-gold standard measures of the construct of interest.	Yes	Yes	See below
a) Convergent and discriminant validity	A measure is correlated with existing validated measures which purport to measure the same construct (convergent validity) and uncorrelated (or less strongly correlated) with measures which purport to measure a different construct (discriminant validity). When using non-gold standards it may be uncertain how far non-agreement with established tests reflects inadequacies on the part of the existing tests and how far it reflects inadequacies in the new measure.	Yes	Yes	ICC, Pearson's or, Spearman's correlation
b) Group differentiation/ hypothesis testing	The measure can discriminate between groups known to be different with regard to the characteristic of interest (group differentiation) or believed to be different with regard to the characteristic of interest (hypothesis testing). In child mental health, examples include the expectation of worse mental health among children in a clinic sample than in a community sample. Another form of hypothesis testing sometimes used is responsiveness, which refers to the fact that the measure should show improvement after a treatment of known efficacy. This forms a subset of methods which test hypotheses relating to predicting future events.	Yes	Yes	Sensitivity, specificity, ROC analyses
c) Factor analysis	Whether the pattern of interrelation between items corresponds to the hypothesised factor structure of the measure. This includes items which propose to be measuring the same subscale being more closely related to each other than to items from different nominal subscales. It may be difficult to apply these techniques to diagnostic interviews if extensive skip rules mean that many items were not asked of all individuals.	Yes	Difficult if skip rules used.	Exploratory and confirmatory factor analyses

Adapted from Lamping *et al.* [560-561], Hilari *et al.* [562] and Smith *et al.* [563]

13.1.2 Measures of agreement

Below I describe five different measures of agreement which I either use myself in this thesis or which I report from previous studies. After describing these measures, I then discuss briefly some of the main considerations in choosing the appropriate measure for a particular analysis.

Intraclass Correlations

Intraclass correlations (ICCs) form a family of measures, which can be calculated within the framework of analysis of variance (ANOVA) models [564]. The intention of an ICC is to partition the total variance between individual measurements into variance between subjects (e.g. between different children) and variance within subjects (e.g. different measurements of the same child). The basic formula for ICCs is

Equation 13.1: General formula for the intraclass correlation coefficient (ICC)

$$\text{ICC} = \frac{\text{Between-subject variance}}{(\text{Between-subject variance} + \text{within-subject variance})}$$

The ICC therefore varies from 0 to 1, with higher values indicating that within-subject variance is small relative to between-subject variance. To say that within-subject variance is comparatively small is equivalent to saying that within-subject correlation is comparatively high; hence the name ‘intraclass correlation’.

There are multiple versions of the ICC, with the correct version depending on three key questions [564]:

1. Whether the raters are different for each individual observed or whether the same raters have rated all individuals in the dataset. If the raters are all different then it is not possible to separate out a between-subject variance component due to the effect of differences between individual raters – instead this is indistinguishable from other sources of between-subject variation. This corresponds to a one-way ANOVA. If the same raters have been used for multiple subjects then one can use a two-way ANOVA to look at rater effects as well as child effects

2. If the same raters have rated multiple subjects, the next question is whether differences between the raters' mean ratings are important. If one wants to look at levels of absolute agreement between raters then differences in mean rating are important and a random-effects model for the rater effect should be used. By contrast if one is merely interested in *consistency* across raters then one can treat differences in means between different raters as a fixed effect.
3. Whether one is using single ratings or mean scores from several ratings.

Pearson's correlation coefficient

Pearson product-moment correlation coefficient is a measure of the linear dependence between two variables, X and Y [496]. The statistic is calculated as the product of the standardised scores of each X and its corresponding Y. These are then summed and divided by the degrees of freedom to give the final value of the coefficient. If the data is taken from a sample, the statistic is calculated as follows:

Equation 13.2: Formula for Pearson's correlation coefficient

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{s_X}\right)\left(\frac{Y_i-\bar{Y}}{s_Y}\right)$$

Where $\bar{X}$ is the mean, ***Sx*** the standard deviation, $\left(\frac{X_i-\bar{X}}{s_X}\right)$ the standardised value of the *i*th observation of X, and (n-1) is the degrees of freedom. If the data are taken from a total population rather than a sample, then the degrees of freedom is n and the population mean and standard deviation are used.

The value of the Pearson's correlation ranges from -1 to +1, with a correlation of +1 indicating a perfect positive linear relationship between variables.

The main assumptions of the Pearson correlation coefficient are

- That X and Y are independent.
- That the X and Y variable are both continuous and measured on an interval scale.
- That the relationship between two variables is linear (rather than curved).

Comparing Pearson correlation coefficients

Fisher's z-transformation provides a method of converting Pearson's correlation coefficients into values with an approximately normal distribution [565]. The formula of Fisher's z-transformation is:

Equation 13.3: Formula for Fisher's z-transformation

$$\text{Fisher's z-transformation} = 0.5 * \ln[(1+r) / (1-r)]$$

Where r is the correlation coefficient. The result is a distribution which is approximately normal and which has a standard error of $\sqrt{[1/(n-3)]}$. Assuming the correlation coefficients are independent, one can then test for equality between two transformed values using standard techniques for comparing two independent, normally distributed random variables. That is, one can calculate the z-statistic as follows:

Equation 13.4: Calculating z-statistics from two independent, normally-distributed random variables.

$$\text{z-statistic} = \frac{\text{Difference in point estimates}}{\text{Pooled standard error of point estimates}} = \frac{|z_1 - z_2|}{\sqrt{[1/(n_1-3) + 1/(n_2-3)]}}$$

Spearman's correlation

Spearman's correlation coefficient is a non-parametric alternative to Pearson's which does not assume the variables are normally distributed and measured on an interval scale [496, 566]. Instead, the values of the data are first converted into ranks before the correlation coefficient is calculated. The equation whereby this is done is identical to that of the Pearson's correlation (Equation 13.2) for untied data, and can also be simplified to the following:

Equation 13.5: Formula for Spearman's correlation coefficient

$$r = 1 - \frac{6 \sum_{i=1}^{n} D_i^2}{n(n^2 - 1)}$$

Where D_i is the difference between the ranks of the ith observation of X_i and Y_i between the two measures. For tied observations, one takes the mean of the ranks associated with the tied observations and the value of the resulting coefficient is therefore slightly different to the Pearson equivalent.

Comparing Spearman correlation coefficients

Spearman's coefficients can be compared using the same Fisher's z-transformation described above in Equation 13.3. This has been shown to function satisfactorily for the purposes of making inferences [567]. It is also at least as robust to type I errors as the two alternative approaches of 1) calculating Pearson coefficients even if the assumptions are not met; or 2) calculating Spearman's coefficients and then using a conversion formula to turn them into Pearson's coefficients.

Kendall tau

Although Spearman's coefficients are non-parametric, they do assume that the size of the intervals between ranks can be quantified in a comparable way across all ranks – for example, that a difference of rank 10 vs. rank 13 is equivalent to 3 units and therefore comparable to a difference between rank 2 and rank 5. Kendall's tau (τ) is an alternative for ordered categorical data which does not involve any assumptions about the size of the intervals between pairs of ranks [566]. Instead it measures monotonic agreement, looking at agreement about the position of items relative to each other. For example, a set of children may be rated by both their teachers and their parents. The parents and teachers are said to 'agree' with respect to children A and B if the rank of child A by parent report is higher than the rank of child B by parent report, and if the rank of child A by teacher report is also higher than the rank of child B by teacher report. By contrast, if the rank of child A by teacher report were lower than the rank of child B by teacher report, then the two informants would disagree.

Equation 13.6: Formula for Kendall's Tau

$$\text{Tau} = \frac{P - Q}{\frac{1}{2}\, n\,(n-1)}$$

Where P = the number of items (e.g. children) ranked in the same order by the raters (e.g. parents and teachers) and Q = the number of items in which the rankings are in the opposite order.

Cohen's weighted chance-corrected kappa (κ) statistics for agreement

Cohen's kappa measures the agreement between two raters who independently classify N items into a fixed number of mutually exclusive categories. It then adjusts this for the number of agreements which would be expected by chance alone based on the marginal frequencies with which each rater chooses each category. In general, the formula for chance-corrected agreement can be presented in any of the following forms [496, 566]:

Equation 13.7: Alternative expressions of the formula for chance-corrected agreement

$$\text{Agreement} = 1 - \frac{\text{observed disagreement}}{\text{\% expected chance disagreement}}$$

$$= 1 - \frac{(1 - \text{\% observed agreement})}{(1 - \text{\% agreement chance expected})}$$

$$= \frac{(1 - \text{\% expected chance agreement}) - (1 - \text{\% observed agreement})}{(1 - \text{\% expected chance agreement})}$$

$$= \frac{\text{\% observed agreement} - \text{\% expected chance agreement}}{(1 - \text{\% expected chance agreement})}$$

One version of the formula for the chance-corrected kappa statistic is therefore:

Equation 13.8: Formula for the chance-corrected Kappa

$$\text{kappa} = \frac{p_0 - p_e}{1 - p_e}$$

Where p_0 equals the observed probability of agreement between raters (i.e. the proportion of instances in which the raters agreed), and p_e equals the expected probability of agreement between raters by chance, as calculated from the marginal probabilities. The kappa therefore takes a maximum of 1 for perfect agreement, 0 for agreement no better than chance and is negative for agreement worse than chance.

The kappa can be used for binary, ordered categorical or unordered categorical data. For binary and unordered categorical data, exact agreements are counted in the 'observed agreement' total. This is the unweighted kappa. For categorical data, one can assign weights to instances of disagreement such that greater weight is given to cases of near-agreement. In a table with g categories over all, a cell in row i, column j is assigned the weight:

Equation 13.9: Weights for weighted chance-corrected Kappa

$$w_{ij} = \frac{1 - |i-j|}{g-1}$$

Considerations in choosing appropriate measures of agreement

Two raters vs. more than two raters

Pearson's, Spearman's, Kendall's tau and chance-corrected kappa all measure agreement between two different raters. By contrast, the ICC can be applied to two raters but can also be extended to three or more raters. This is because ICC uses a framework of partitioning sources of variance, rather than calculating measures of distance between paired observations.

Source of the data and nature of the data

As described above (p.358), two of the key questions when choosing the appropriate ICC are whether the raters are the same across all individuals and whether the scores for comparison are individual ratings or means. This will clearly vary between studies. In the case of B-CAMHS, the raters for different children are not the same – for example, it is not the *same* teacher who rates each of the children but rather a different teacher each time. This thesis (and other analyses of B-CAMHS) also use individual SDQ scores rather than mean ratings. This corresponds to a one-way ANOVA model for individual values, and to the form of ICC referred to as ICC(1,1).

Relative agreement vs. absolute agreement?

The third question in deciding the appropriate ICC was whether differences between the raters mean ratings are important. That is, if one rater gives children lower scores on average than another rather, does this matter? This can be thought of as corresponding to the question of whether absolute agreement matters as well as relative agreement [496].

Pearson's, Spearman's and Kendall's tau coefficients measure relative agreement only; none of these three measures is affected if (say) one adds 10 to each score from one particular rater. By contrast, chance-corrected kappa's only measure absolute agreement between categories. ICCs without shared raters (including the ICC(1,1) appropriate for B-CAMHS) are influenced by both relative agreement and absolute agreement. If raters are shared, one has the additional option of choosing to partition out mean differences between particular raters and so examine only relative difference.

Clearly whether relative or absolute agreement is important depends on both the nature of the data and the purpose of one's analyses. For example, if one is measuring agreement between psychiatrists in assessing the severity of a condition with a view to deciding whether to initiate treatment, then it is clearly important if some psychiatrists systematically give higher scores to the same child than their colleagues. By contrast, if one is measuring how far different types of informant (e.g. parents, teachers and children) agree about the mental health of a particular child then it may not be important if (say) children systematically give higher means scores than parents. What is instead of interest is the relative agreement between these individuals.

Distribution of the data

Finally, the distribution of the data is relevant to deciding which measures are appropriate. As is often the case, there is a trade-off between the amount of information one retains from the data and the strength of the assumptions one is prepared to make. As illustrated in Table 13.3, Pearson's coefficient is more powerful than Spearman's because it retains the original data, but as described above, it is a parametric measure which requires stronger assumptions about a linear relationship between normally variables. Spearman's coefficient, in turn, retains more of the original data than Kendall's tau but also requires the additional assumption that it is meaningful to quantify the difference between ranks rather than just judging them to be 'larger' or 'smaller'

Table 13.3: Data transformation for calculating Pearson's, Spearman's and Kendall's tau

Raw data	10	11	15	19	22	40
Data used for Pearson's	10	11	15	19	22	40
Data used for Spearman's	1	2	3	4	5	6
Data used for Kendall's Tau	Smaller	Smaller	[reference]	Larger	Larger	Larger

13.1.3 Internal consistency

Cronbach alpha (α)

The formula for Cronbach alpha is:

Equation 13.10: Formula for Cronbach alpha

$$\alpha = \frac{N}{N - 1} \left(\frac{1 - \Sigma^{N}_{i=1} \sigma^2_j}{\sigma^2_X} \right)$$

or equivalently

$$\alpha = \frac{N}{N - 1} \; \frac{\Sigma\Sigma^{N}_{i \neq k} \sigma_{jk}}{\sigma^2_X}$$

where N is the number of items, σ^2_j is the variance of each item score X_j, σ_{jk} is the covariance between item scores X_i and X_j and σ^2_X is the variance of the observed total test scores (i.e. the sum of all the items). The value of Cronbach alpha ranges from 0 to 1. It can be calculated for as few as two items but, because of the N/N-1 part of the equation, tends to increase as the number of items in the scale increases.

As the sum of the inter-item covariances divided by the total variance (formulation two of Equation 13.10), Cronbach alpha gives the average "interrelatedness" of items, assuming that none of the covariances are negative. Cronbach alpha coefficient is widely used as a measure of the internal consistency [245, 496] with values in the range of 0.7 to 0.9 generally considered 'acceptable'. The use of Cronbach alpha in this way has, however, been widely critiqued in the psychometrics literature [e.g. 559, 568]. There has also been critique of the widespread use of Cronbach alpha as lower bound to reliability in preference to alternative methods such as the greatest lower bound [569].

13.1.4 Criterion validity and construct validity

Sensitivity and specificity

The sensitivity of a test is the proportion of genuine 'cases' correctly identified (a/(a+c) in Table 13.4). The specificity of a test is the proportion of genuine 'non-cases' correctly identified (d/(b+d) in Table 13.4). In a perfect test, both values will be equal to 1. In practice, the two typically trade off against each other for any given method of testing, with gains in sensitivity (i.e. a higher proportion of cases correctly classified as cases) coming at the expense of reduced specificity (i.e. a higher proportion of non-cases incorrectly classified as cases).

Table 13.4: Calculation of sensitivity and specificity

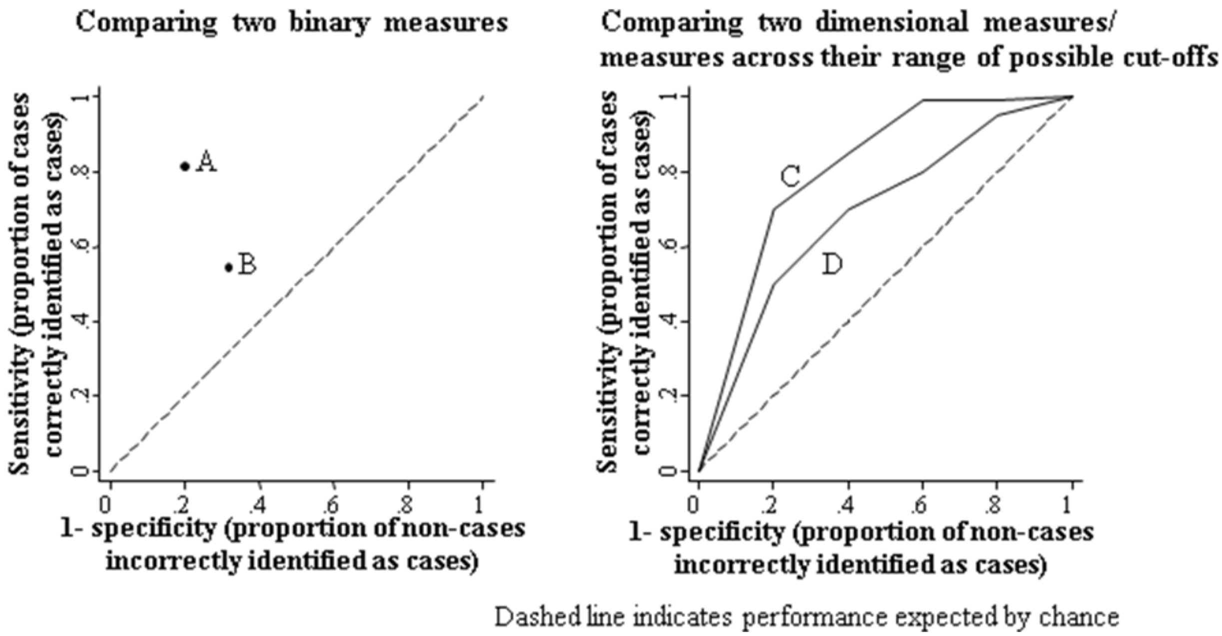
		Gold standard		
		Case	**Non-case**	TOTAL
Measure	**Case**	a	b	a+b
	Non-case	c	d	c+d
	TOTAL	a+c	b+d	

Receiver Operating Characteristics (ROC) analyses

Receiver Operating characteristic (ROC) curves represent a way to visualise and compare the ability of a test to correctly differentiate between groups [570]. These can be groups defined according to a gold standard in which case this is a measure of criterion validity. Alternatively they can be groups defined as high- and low-risk according to some other characteristic (e.g. psychiatric clinic attenders) in which case this is a measure of group differentiation.

As illustrated in Figure 13.1, ROC curves plot sensitivity (a/(a+c) in Table 13.4) against one minus the specificity (one minus b/(b+d) in Table 13.4). The former corresponds to the proportion of cases correctly identified as cases, the latter to the proportion of non-cases incorrectly identified as cases. The diagonal line represents performance expected by chance, with anything 'northwest' of this line indicating performance better than chance and the top left corner representing a perfect test. Sometimes, as in the left-hand graph, ROC analyses can be used to compare different tests; in this case, A is superior to test B because it produces more true positives and fewer false positives. ROC analyses can also be used to visualise the performance of different cut-offs from the same test (for example, line C in the right-hand graph of Figure 13.1), and/or to compare different tests across their range (for example, comparing line C and line D). The latter involves a comparison of the Area Under the Curve (AUC) for each measure, with higher values indicating a better test.

Figure 13.1: Schematic illustration of ROC curves



Multitrait-multimethod analyses

Multitrait-multimethod (MTMM) analyses are a method for assessing the construct validity of a set of measures [494-495]. As illustrated in Table 8.1, MTMM are based on a correlation matrix of multiple traits (e.g. the proposed SDQ subscales) as measured by multiple methods (e.g. by parent and teacher). MTMM analyses also typically present the interrelationship between different traits measured by the same method, and the test-retest reliability of each trait. Together, this provides evidence on several aspects of convergent and discriminant validity[41]:

1) Test-retest reliability for each trait ([a] cells in Table 8.1).

- Good convergent validity requires high correlation between two measurements of the same trait by the same method after an appropriate time interval. This could not be assessed in B-CAMHS because the necessary retest was not carried out.

[41] As described in Table 13.2, convergent validity requires highly correlation with other measures of the same or similar traits while discriminant validity requires little or no correlation with measures of different traits.

2) Within-method correlation of different traits ([b] cells, in the heterotrait-monomethod triangles).

- Good construct validity requires within-method correlations between traits which are consistent with *a priori* hypotheses. For example, if prosocial behaviour is hypothesised to be more closely related to externalising than internalising problems, then the prosocial-externalising coefficients should be larger than the prosocial-internalising coefficients within all informants. If this were not observed then it would suggest that the method in question was not providing a valid measure of the hypothesised constructs.
- The heterotrait-monomethod triangles therefore provide a means of hypothesis testing regarding the expected relationship between different subscales. They also address one aspect of discriminant validity, namely the requirement that the within-method inter-trait correlations not be too high. This is because correlations which are high to the point of collinearity indicate that the measures are in fact measuring the same traits.

3) Between-method correlation of the same trait ([c] cells or validity coefficients, shaded grey).

- Good convergent validity requires high correlation between measures of the same trait assessed by means of different methods (e.g. parent externalising score and teacher externalising score).
- In addition, the magnitude of the [c]-cells should ideally be larger than the [b] cells. If not then this indicates that the 'method factor' (i.e. the informant) is a more powerful determinant of a child's score than the 'trait factor' (i.e. the child's mental health).

4) Between-method correlation of different traits ([d] cells).

- Good discriminant validity requires that the correlation between different traits be substantially lower than the correlation between the same traits. So, for example, the parent externalising score should be substantially less highly correlated with the teacher internalising and prosocial scores than with the teacher externalising score.

Table 13.5: Schematic representation of an MTMM analysis for the parent and teacher SDQ

		Parent			Teacher		
		Internalising	**Externalising**	**Prosocial**	**Internalising**	**Externalising**	**Prosocial**
Parent	**Internalising**	*[a]*					
	Externalising	*[b]*	*[a]*				
	Prosocial	*[b]*	*[b]*	*[a]*			
Teacher	**Internalising**	*[c]*	*[d]*	*[d]*	*[a]*		
	Externalising	*[d]*	*[c]*	*[d]*	*[b]*	*[a]*	
	Prosocial	*[d]*	*[d]*	*[c]*	*[b]*	*[b]*	*[a]*

Cells labelled [a] show the test retest reliability coefficients for each subscale for each informant. Cells labelled [b] show agreement between different subscales reported by the same informant, and form heterotrait-monomethod triangles. Cells labelled [c] are the validity coefficients, and show correlations between the same subscales reported by different informants. Cells labelled [d] show agreement between different subscales reported by different informants. Together the [c] and [d] cells form a heterotrait block.

13.1.5 Factor analysis

Factor analysis provide a family of techniques which derive a small set of unobserved ('latent') factors which account for the covariance between a larger set of observed ('manifest') variables [490, 571]. For example, the items on a questionnaire may be the manifest variables, and the proposed construct(s) which they measure would be the latent variable(s) or factor(s). Factor analyses can either be conducted in a data-driven way (exploratory factor analysis or EFA) or in a hypothesis-driven way (confirmatory factor analysis or CFA). I use both techniques in this thesis, with confirmatory factor analyses playing a particularly crucial role.

Exploratory factor analysis (EFA)

Exploratory factor analysis imposes no constraints upon the pattern of relationships seen between the underlying factor and the manifest variables [490, 571]. Instead each latent factor is assumed to affect all manifest variables as follows:

Equation 13.11: Model for exploratory factor analysis

$$x_1 = b_{11}f_1 + b_{12}f_2 + b_{13}f_3 + \ldots + \mu_1 + e_1$$

$$x_2 = b_{21}f_1 + b_{22}f_2 + b_{23}f_3 + \ldots + \mu_2 + e_2$$

Where x_1 to x_p are the manifest variables and f_1, f_2, f_3… are the latent factors. μ_i is the mean of x_i and e_i is the residual. The coefficients of the factors, b_{ij}, are constants estimating how much each manifest variable is influenced by each factor – these are called the factor loadings. Where the standardised x's are used, the factor loading b_{ij} is equal to the *i*th manifest variable and the *j*th factor. The variance in x_i which is explained by the ith factor model is the communality (see also discussion of sample size below, p.375).

Often a rotation of the factor loadings is used in order to maximise high loadings, minimise small ones and therefore clarify the model structure. Such rotations can be either orthogonal (in which the latent factor are assumed to be uncorrelated) or oblique/non-orthogonal (in which the latent factors are allowed to correlate). The Geomin rotation is one oblique rotation which is indicated by simulation studies to be particularly suitable when little is known about the underlying structure, performing well for both simple models where all items load primarily onto a single factor and also in more complex situations [572]. I therefore use this rotation in this PhD.

Exploratory factor analysis vs. principal component analysis

Principal component analysis (PCA) is sometimes used as an alternative means of exploring the internal structure of a set of manifest variables. It represents an orthogonal linear transformation of a set of data items as follows:

Equation 13.12: Model for principal component analysis

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p$$
$$y_2 = a_{21}x_1 + a2_2x_2 + \ldots + a_{2p}x_p$$

where y_1 is the first 'principal component', y_2 the second and so on. Each principal component is calculated so as to be orthogonal to the previous components (although rotations can then be used to accommodate oblique solutions). The values of a_{1j} in the first linear combination are calculated so as to maximise the variance of y_1, the values of a_{2j} to maximise the variance of y_2 and so on.

As such, while EFA fits a model which assumes the latent factors predict the manifest outcome variable, PCA fits a model in which the manifest variables predict the (latent) principal component. PCA also differs in ignoring the possibility of measurement error (the 'e_i' term in Equation 13.12). For these reasons, PCA is not ideally suited to investigations of internal structure and can lead to poor estimates of factor loadings in small samples [573-574]. Nevertheless the use of PCA remains widespread for this purpose – and, indeed, the only 'factor analyses' reported to date from B-CAMHS [448] in fact used PCA (Robert Goodman – personal communication).

In addition, even when using EFA, many researchers use Eigenvalues [575-576] to decide upon the number of factors to specify. Each Eigenvalue is calculated as equal to the variance of one of the principal components such that, if the standardised values of the x's are used (and given a total of *p* manifest variables), then

Equation 13.13: Relation of Eigenvalues to variance of the manifest variables and principal components

$$\Sigma \mathrm{var}(x_i) = \Sigma \mathrm{var}(y_i) = \Sigma \lambda_i = p$$

The contribution of each Eigenvalue to the total observed variance is therefore λ_i/p. Typically, '1' is used as a cut-off for determining the number of factors, such that one only includes factors which explain a greater proportion of the variance that would be explained by a single factor (where $\lambda_1 = 1$).

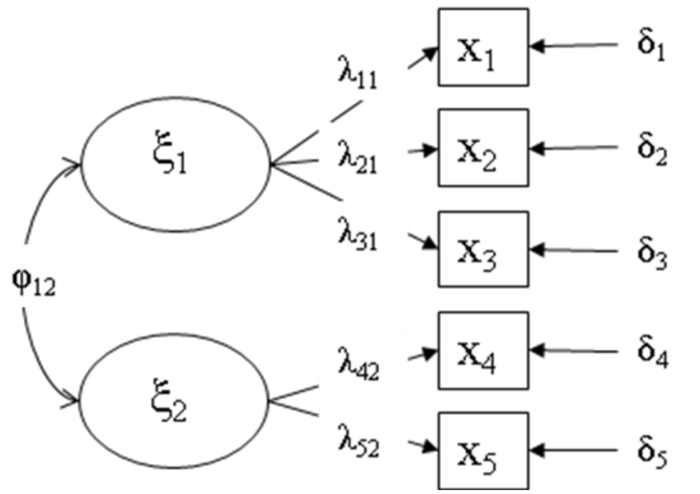
Confirmatory factor analyses

Both principal component and principal factor analyses are exploratory techniques, primarily useful in suggesting plausible factor structures when these are not known. When a hypothesised factor structure exists, it is more appropriate to use a model-based framework such as confirmatory factor analysis (CFA). This allows hypothesis-driven statistical testing of whether a proposed factor structure fits the data, and formal comparisons of the relative fit of competing factor structures [490]. Fitting CFA therefore involves three stages: model specification, model estimation and evaluation of goodness-of-fit.

Model specification

To fit a CFA one must specify the hypothesised number of latent factors, their relation to the manifest variables and their relation to each other. Conventionally, squares are used to represent the observed manifest variables and circles the latent variables. Single-headed arrows indicate a causal relation and double-headed arrows show covariance between variables [571]. So, for example, Figure 13.2 starts with five observed variables (x_1 to x_5), each with error variance δ_i. These manifest variables are hypothesised to be influenced by two latent traits (ξ_1 and ξ_2), which have a covariance of φ_{21}. The λ_{ij}'s are the factor loadings (equivalent to the b_{ij}'s above). For a CFA model to be identified some constraint has to be placed on the factor loadings; typically, these are either to set one of the values of the factor loadings to 1 or to set their variance to 1.

Figure 13.2: Illustration of model specification in a first order CFA



Model estimation

The equation of each manifest variable in Figure 13.2 can therefore be written as follows:

$x_1 = \lambda_{11}\,\xi_1 + \delta_1$

$x_2 = \lambda_{21}\,\xi_1 + \delta_2$ etc.

This model specification clearly has many similarities with linear regression (discussed more fully in Section 13.2), but with the important difference that in factor analysis the latent variables (ξ) are not observed. As such, the parameters λ_{ij} are estimated by calculating the observed covariance matrix between the X's and choosing parameters for λ_{ij} in order to give a predicted covariance matrix which is as similar as possible. This can be done using maximum likelihood estimation (MLE) under the assumption of multivariate normality among the manifest variables. For categorical manifest variables a weighted least squares (WLS) approach is preferable, and this performs better at achieving nominal rejection rates when significance testing (i.e. 5% of models rejected at $p<0.05$) and in estimating the magnitude of loadings [577]. When using binary data, it is also possible to fit the CFA using multivariate probit analysis rather than a multivariate normal model [491], with extensions being available for ordinal data [492].

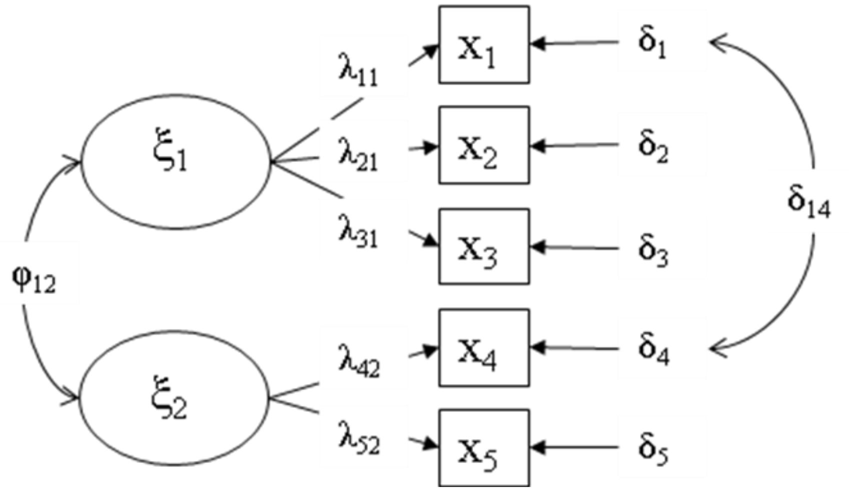
Assessing goodness of fit

There exists considerable controversy regarding which is the best of the available statistics for assessing model fit in a CFA, and what cut-offs should be used [490, 493]. Chi-squared test statistics are problematic as they are sensitive the sample size, being increasingly likely to reject appropriate models as the sample size increases. I therefore follow common practice in reporting more than one of the alternative measures of fit of fit which have been developed, these being the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI) and the Root Mean Square Error of Approximation (RMSEA).

To consider a model as showing acceptable fit, I required the CFI to be >0.90 and ideally >0.95; the TLI to be >0.90 and ideally >0.95; and the RMSEA to be <0.08 and ideally <0.05 [490]. In some cases, it may be possible to improve the fit of a model somewhat by allowing the error of different observations to be correlated, as shown by the term δ_{14} in Figure 13.3. Such correlation may occur because of shared measurement error due to a rater bias (e.g. both x_1 and x_4 are reported by the same informant) or because of a bias across time (e.g. x_1 and x_4 represent the same variable measured at two time points). Adding this error covariance to the model reduces the unexplained variance in the model and can therefore improve model fit without changing the substantive hypothesis being tested.

Where the number of manifest variables is large one may achieve good overall model fit even if one has misspecified the model with regard to one particular item. It is therefore also advisable to examine the individual standardised loadings (λ_{ij}) of each manifest variable onto its latent factor(s) to check that these are of reasonable magnitude. This is usually operationalised as >0.4 or above.

Figure 13.3: First order CFA with correlation of residual variances



Multi-group confirmatory factor analyses

Once one has demonstrated the good fit of a model in the pooled sample, one can then go on to fit a multi-group CFA which tests the hypothesis of measurement invariance across different subgroups of the total population. Measurement invariance refers to the situation in which the relationship between manifest and latent variables is the same across all populations – that is, when all manifest variables have equivalent measurement parameters (thresholds, factor loadings and standard errors) across all groups [578-580]. For example, let η denote a given fixed level on the latent trait (ξ) that predicts the manifest variable x. If one is comparing across groups designated by the grouping variable v then measurement invariance implies that, as shown in Equation 13.14, the influence of an underlying factor upon the observed value of a variable should be the same regardless of their group membership.

Equation 13.14: Measurement invariance across groups

$$E(x \mid \eta, v) = E(x \mid \eta)$$

As such, measurement invariance does not imply that the mean level of the underlying trait (ξ) is identical across groups. It does, however, imply that at any given level (η) of that trait, the value observed in the individual does not vary across groups. Measurement invariance therefore indicates that the same underlying constructs exist across groups, that all items are functioning similarly, and that it is therefore meaningful to interpret group differences in the test scores as differences in the underlying constructs. Measurement invariance can be demonstrated by showing that a multi-group CFA shows good fit to the data, with good fit defined in the same way as for a global model (i.e. CFI >0.90 and ideally >0.95; TLI >0.90 and ideally >0.95; and RMSEA <0.08 and ideally <0.05)

Sample size and factor analyses

Inadequate sample size may lead to instability of estimates in both EFA and CFA. There is some controversy regarding the best way to judge adequate sample size for factor analyses. Various rules of thumb are commonly cited such as an "N:p ratio" rule (e.g. 'five subjects per item') or an "absolute N" rule (e.g. minimum of 200 subjects), although the content of these rules varies quite considerably [reviewed in 581]. Drawing on theoretical and empirical data, however, MacCallum *et al.* [581-582] argue that it is fundamentally incorrect to assume that the minimum necessary sample size is invariant across studies. Instead, the quality of factor solutions is increased by a number of features of the analysis under consideration, including:

- A larger absolute sample size.
- A smaller number of hypothesised factors.
- High communality (>0.5) of the manifest variables; that is, a high portion of variance in that item is explained by the common factor. This is represented by the R-squared value for each item.
- Greater overdetermination of the factors; that is, the extent to which each factor is well-defined by a set of manifest variables. In general, a small number of factors defined by a large number of indicators will show greater overdetermination.

There are also important interactive effects between these influences, such that absolute sample size and overdetermination are particularly important if communalities are low.

In the models fitted in the pooled B-CAMHS sample presented in Section 8.1 Chapter 8, communality was generally low, being under 0.5 for half to two-thirds of the items. Under such circumstances, MacCallum *et al.* advise that if "there is high overdetermination of factors (e.g. six or seven indicators per factor and a rather small number of factors), one can still achieve good recovery of population factors, but larger samples are required – probably well over 100” [581, p.96]. My assessment of measurement invariance between Indians and Whites in Section 8.2 Chapter 8 tests a model with two factors (internalising and externalising), each with 10 items. This therefore meets the overdetermination condition. As such, the sample sizes for Indians for parents (N=389) and teachers (N=306) are likely to be adequate. The child sample size of 184 is perhaps somewhat small, however, meaning that the results for the child SDQ should be treated with some caution.

Missing data

In both EFA and CFA analyses using the WLSMV estimator, MPlus deals with individuals who have partial data using pairwise present methods. That is, it estimates parameters using information from pairs of variables and uses all individuals with observations for that pair. Individuals with partial data are therefore retained in the analyses and their information is used for all pairwise analyses for which it is present.

13.2 Regression analyses

Regression analyses provide a large family of models which describe the relationship between two or more variables, with some being outcome (dependent) variables and others explanatory (independent) variables. There can be many motivations for this kind of analysis – among the most common are:

1. Estimating the independent effect of a particular explanatory variable of interest.
2. Identifying the set of explanatory variables most important in predicting the outcome variable.
3. Predicting an outcome variable.

In this thesis I use regression many times, and particularly for the first and second of these purposes. Below I describe in more detail the statistical underpinning and accompanying assumptions of the regression techniques which I employ in this thesis.

13.2.1 Linear regression

Linear regression models are the simplest regression models, and take the general form shown in Equation 13.15.

Equation 13.15: Formula for linear regression

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots \beta_n x_{ni} + e_i \qquad \text{with } i = 1, 2 \ldots n,\ e_i \mid X \sim IID(0, \varphi^2)$$

With the corresponding prediction equation

$$E(y_i) = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots \beta_n x_{ni}$$

In this model, $E(Y_i)$ is the expected value of the random variable y for person *i* based on the values of x_{1i}, x_{2i} ...x_{ni} – that is, the conditional expectation of the random variable . β_1 and β_2 are partial regression coefficients and in continuous variables measure the effect per unit increase in one explanatory variable controlled for the other. For example, β_1 is the expected increase in y per unit increase in x_1, holding $x_2 \ldots x_{ni}$ constant. In binary/categorical variables the partial regression coefficients measure the effect of taking the value of ‘1’ for the level of the variable in question as opposed to the baseline group.

The constant α is the expected value of y_i where x_1, x_2 etc are all zero. Finally, because most models will not estimate the outcome perfectly, there is a term e_i for the error or residual – that is, the discrepancy between the predicted value $E(Y_i)$ and the observed value y_i. A fundamental assumption of linear regression is that these residuals are independent of each other, and have a constant variance for each value of X.

Linear regression models can be straightforwardly estimated using ordinary least squares method, where α and the β's are chosen so as to minimise the sum of squared residuals (e_i's) of the y variables. In this instance, ordinary least squares regression is equivalent to using a maximum likelihood method.

Assumptions of linear regression

- The relationship between the outcome variable (y) and explanatory variable(s) (x_1, x_2 ...x_n) is linear.
- The observations of the outcome variable are independent of each other (so not, for example, time series data).
- The errors (e_i's) are independent, with a mean of zero and a constant variance (homoskedasticity) for each value of X. That is, for $X = x_i$, $e_i \sim IID(0, \varphi^2)$ [IID = identically independently distributed].
- For Maximum Likelihood estimation methods, it is also necessary that the distribution of the residuals is approximately normal; i.e. $X = x_i$, $e_i \sim NID(0, \varphi^2)$ [NID = normally independently distributed] This is not essential for ordinary least squares regression in linear regression unless the sample is small [583]. All linear regression models presented in this thesis do require this assumption, however, as they use pseudomaximum likelihood estimation to adjust for complex survey design (see Section 13.3.2).

Examining the assumptions of linear regression[42]

- **Linearity (1)**: Plot the outcome variable against each of the explanatory variables; the relationship should be approximately linear. It is, however, not easy to examine

[42] These methods of testing assumptions also apply to other generalised linear models, except that when using the outcome variable one first transforms it by the link function and that for the 'expected values' one uses the linear predictor (see below, p. 382).

graphically whether the effect of an explanatory variable is linear in a model which simultaneously adjusts for multiple other explanatory variables. One possibility in this situation is to enter the explanatory variable of interest into the model as a categorical variable, and then do a likelihood ratio test to see if this fits better than a model in which it is entered as a linear term.

- **Linearity (2):** Plot the residuals against the expected values: if the model appropriate, these should show random scatter around zero.
- **Constant variance of the errors**: Plot the residuals against the explanatory variable; check there is no tendency for the scatter to increase or decrease at higher values.
- **Normality of the errors**: Normal plots of the standardized residuals
- **Identify data points with large influence**: The leverage of a particular individual's data refers to how great an effect their data point has upon the value of the regression slope. These can be identified by plotting leverage against the squared normalised residuals or quantified by statistics such as the Cook's distance. Points which have high leverage and are outliers (i.e. have a large residual) are highly influential. The data on which they are based should be checked for errors and, if correct, sensitivity analyses excluding these points should be performed.

Violation of assumptions: transformations using logs and zero-skew logs

If the assumptions of the linear regression model are not met, one common approach is to transform the data using a transformation such as taking (natural) logarithms. This may both stabilise the variance of residuals and improve their normality. If the data is positively skewed one can also extend the taking of logarithms to taking 'zero-skew' logarithms. This takes logarithms after adding or subtracting a certain amount determined empirically from the data such that it improves the spacing of the points to remove any skew. The zero-skew logs for the parent, teacher and child SDQ are presented in Table 13.6.

Table 13.6: Formula for zero-skew logs of parent, teacher and child SDQs

Raw outcome	**Log-transformed outcome**	**Zero-skew logged outcome**
Parent TDS	ln(parent TDS)	ln(parent TDS + 4.33)
Teacher TDS	ln(teacher TDS)	ln(teacher TDS + 2.35)
Child TDS	ln(child TDS)	ln(child TDS + 14.19)

A further benefit of zero-skew logs compared to natural logarithms is that they can be used on variables like the SDQ which take the value 0 (whereas for ln(0) this is minus infinity). The disadvantage with zero-skew logs, and indeed with transformations more generally, is that they make the interpretation of the outcome variable more complicated – i.e. 'expected change per unit increase in x in the value of [zero-skew] log of y' rather than 'expected change per unit increase in x in the value of y'. In addition, the skew of a variable is likely to vary somewhat between different samples, and therefore the exact equation of the zero-skew logs will vary. Clearly back-transforming one's final results is a possibility, but nevertheless the use of zero-skew logs as outcomes for multiple intermediate models may hinder the comparison of results between studies.

Standardised regression coefficients

In addition to using transformations in order to meet model assumptions, it is also common to transform both outcome and explanatory variables by dividing them by their estimated standard deviation. One benefit of this is to facilitate interpretation of variables measured using non-intuitive units such as scores on a questionnaire. For example, to someone unfamiliar with the SDQ, knowing that an intervention improves children's scores by 'one SDQ point' may not tell them as much as knowing it improves children's scores by 'a quarter of a standard deviation'. Taking standardised residuals also facilitates comparisons across models which use different measures to capture the same construct. For example, knowing that one intervention improves children's SDQ scores by 'one point' and that another intervention improves children's Rutter scores by 'two points' is less useful in judging which is more effective than knowing (say) that the first improves children's SDQ scores by a quarter of a standard deviation and the second improves their Rutter scores by a third of a standard deviation

Comparing independent regression coefficients

It is possible to compare the regression coefficients from the same model fitted in independent samples using the extension of the Z-test for the comparison of two means shown in Equation 13.16 [584-585]. This can also be used to compare two independent standardised regression coefficients.

Equation 13.16: Comparison of regression coefficients from models fitted in independent samples

$$Z = \frac{b_1 - b_2}{\sqrt{(SEb_1^2 + SEb_2^2)}}$$

Where b_1 and b_2 are the regression coefficients, and SEb_1 and SEb_2 are their standard errors.

13.2.2 Binary, multinomial and ordered logistic regression

Generalised linear models

Linear regression is the simplest example of a generalised linear model (GLM) [583]. Generalised linear models provide a framework for thinking about many sorts of regression models. A GLM has three components:

1. **A linear predictor** (η), which is a linear (additive) combination of explanatory variables multiplied by unknown regression coefficients: For the *i*th individual, this takes the form: $\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} ...$ [This is also sometimes written as in matrix notation as $\underline{\mathbf{X}}\beta$, where $\underline{\mathbf{X}}$ is the matrix of observed values of 'x' for each individual and β is a vector of the parameters which are applied to these]
2. **A response distribution function**: Again, the expected value of the outcome (E(Y)) is assumed to be based upon independent observations of Y_i, $(i = 1 \ldots n)$, arising from a distribution function from the exponential family. To fit the model, one must specify which distribution function is being used.
3. A **Link function** relates the linear predictor part of the model, η, to E(Y, as shown in Equation 13.17). The choice of link function depends on the response distribution of Y. The link function is selected so as to provide a suitable scale for the effects of the explanatory variables on the linear predictor, with this typically being selected such that $E(Y_i)$ is transformed to range from $-\infty$ to $+\infty$.

Equation 13.17: Formula for a GLM

$$g(E(Y_i)) = \eta_i \qquad \text{[or alternatively } E(Y_i) = g^{-1}(\eta_i)\text{]}$$

Linear regression models are the simplest regression models, in that the response distribution of the Y is normal and the link function is the identity function; i.e. $E(Y_i)= \eta_i$ as shown previously in Equation 13.15.

Regression for binary outcomes: logistic regression

If the outcome is binary then the distribution function typically chosen to describe it is the binomial distribution. The most common link function used is the logit link (Equation 13.18), which gives rise to logistic regression [586]. In Equation 13.18, p is the probability that the binary outcome variable takes the value 1 (probability of 'disease'). This takes the place of continuous outcome 'y' in linear regression.

Equation 13.18: Logit link function used for logistic regression

$$\text{logit}(p) = \ln\ (p / (1-p))$$

The use of the logit function therefore means that η is not equal to the expected value of the outcome, as in linear regression, but rather equal to the natural logarithm (ln) of the odds (p/1-p) for the outcome. As in linear regression, η will itself still be an additive model. However, the convention is to exponentiate the results to give a multiplicative model for the odds of disease, rather than an additive model for the log odds of disease. This is illustrated in Equation 13.19 for a simple logistic model with only a single binary explanatory variable, but can be generalised to multiple explanatory variables which may have more than two levels. In the case of a continuous variable of a categorical variable, the odds ratio of the exposure (e^{β}) is the odds ratio of disease per unit increase in the explanatory variable. In the case of a categorical variable it is the increase associated with being in a particular level as opposed to the baseline group.

Equation 13.19: Model for logistic regression

$$\text{log odds of } (p_i = 1) = \alpha + \beta x_i$$

$$\text{log odds } (p_{unexposed} = 1) = \alpha \qquad \text{[unexposed: i.e. } x = 0]$$

$$\text{log odds } (p_{exposed} = 1) = \alpha + \beta \qquad \text{[exposed: i.e. } x = 1]$$

$$\text{log odds } (p_{exposed} = 1) = \text{log odds } (p_{unexposed}) + \beta \qquad \text{[i.e. } \beta \text{ is the log odds ratio]}$$

[exponentiate]

$$\text{odds } (p_{exposed} = 1) = \text{odds of } (p_{unexposed}) \times e^{\beta} \qquad \text{[i.e. } e^{\beta} \text{ is the odds ratio]}$$

The values of α and β are usually estimated by maximum likelihood methods, as is the case for all subsequent regression models presented. In other words, the value of the parameters is chosen such that the likelihood function (or equivalently the log-likelihood function) is maximised. This maximum can be obtained by taking the first derivative of the likelihood and findings where this is equal to zero, and making sure that the second derivative is negative.

Regression for categorical outcomes: multinomial logistic regression

The logistic regression model can be extended to accommodate categorical outcome variables [587-588]. One way in which this can be done is through multinomial logistic regression model (also called polytomous logistic regression). This is essentially a series of standard binomial regression models each of which compares the baseline response category to one other level; for example level 2 vs. level 1, then level 3 vs. level 1 etc. As such, if there are k levels of response then the models fits (k-1) independent binary logistic regression models.

Regression for ordered categorical outcomes: Ordered logistic regression

For unordered categorical variables, multinomial logistic regression is necessary. It is not desirable for ordered categorical variables, however, because ignoring the ordered nature of the data reduces power to detect effects. It is therefore better to utilise the ordered nature of the data through extensions of the logistic regression model.

One such extension is the proportional cumulative odds ordered logistic regression model (often called simply the proportional odds model, and also known as the cumulative logit model) [588-589]. If the number of levels of an ordered variable are equal to k, then the proportional cumulative odds refers to the probability of being in a given category or higher (i.e. ≥k) vs. in a lower category (<k). For example, if one has a three point ordered categorical outcome with levels '1', '2' and '3', the odds ratio is the estimate of being in group 2 or 3 vs. group 1, and of being in group 3 vs. group 1 or 2.

This method therefore makes the proportional odds assumption: that is, that the true (population) odds ratio is for being in category ≥k vs. in category <k is the same for all

values of k. In other words, wherever one 'cuts' the data to make a binary variable, the odds of being in the higher vs. the lower category is the same.

One method for assessing the proportional odds assumption is to perform a likelihood ratio test which compares a proportional odds model with the corresponding multinomial model (i.e. in which the odds ratios between levels are not constrained to be equal). In the case of a single explanatory variable these are nested models and so the likelihood ratio is properly valid. For multiple explanatory variables, one can fit partial proportional odds models in which some subset of the explanatory variables are not constrained to proportional. One can then compare the fit of this model with nested, fully proportional odds models [588].

13.2.3 Model selection

Variables in the model

Throughout this thesis, I present forced entry multiple regression models – that is, showing models which include all explanatory variables of potential interest even if their independent effects are not significant. I do this firstly because I believe it is more transparent. Moreover, there are many limitations to methods of model selection such as backwards stepwise elimination of variables. Such stepwise methods may fail to include all variables which do have an influence on the outcome; may include variables which do not have an effect; can produce unstable results in instances of moderate or high collinearity of independent variables; and as a result may often produce suboptimal models [590-591]. Stepwise regression methods can also lead to inflated rate type 1 errors above the nominal rate (e.g. above 5% for "p=0.05") in a way which is not true of forced entry methods [590, 592-593].

The main advantage of stepwise regression methods (or other related methods of model selection) is to achieve a more parsimonious model (i.e. with fewer explanatory variables). This is, however, primarily of value when one is seeking to develop a prediction equation for an outcome that can then be used routinely. This is never the case in this thesis; rather my rationale in fitting multiple regression models is always either to estimate the independent effect of a particular variable (in particular, White vs. Indian ethnicity) or to see which of a set of explanatory variables has the strongest effect on the outcome (e.g. to examine the construct validity of the SDQ subscales relative to the DAWBA). For these purposes, the inclusion of some variables which do not predict the outcome creates more cumbersome models and reduces the degrees of freedom. It is not, however, expected to alter substantially the point estimates of the variables in the model or to obscure which variables have the strongest effects.

Modelling variables

The likelihood ratio test statistic, Λ, is calculated as the ratio of the likelihood of two models, one of which is nested inside the other (i.e. is a special case of that model). For large sample sizes, the test statistic of $-2\ln(\Lambda)$ has a χ^2 distributed with degrees of freedom equal to the difference in the number of degrees of freedom of the two models being compared. The likelihood ratio test statistic can therefore be used to examine whether there is evidence that the simpler nested model provides inferior fit (i.e. a lower maximum likelihood) than the broader model. If not, the nested model may be preferred on the grounds of parsimony.

At various points in this PhD, I use likelihood ratio tests to distinguish between different ways of modelling variables. In particular, I use likelihood ratio tests to compare:

1. Models in which a variable is entered as a categorical term vs. the nested model in which a variable is entered as a linear term
2. Non-proportional odds model in which the odds ratio for a parameter of interest is allowed to differ for different levels of the outcome vs. the nested proportional odds models in which the odds ratio for a parameter of interest is constrained to be the same across all levels of the outcome (proportional odds assumption).

13.3 Complex survey design

13.3.1 General principles

As described in Section 5.2.1 Chapter 5, B-CAMHS employed a stratified sampling design in which the primary sampling units were postal sectors and in which a fixed number of individuals were selected within each sector. Probability weights were calculated to adjust for over- or under-sampling of strata in different countries in Britain, and for differential non-response by region, age and sex. Below I describe these concepts in more detail, before explaining in Section 13.3.2 how I dealt with the B-CAMHS sampling design in my analysis.

<u>Stratification</u>

In many surveys, the total population is first divided up into pre-specified¸ non-overlapping 'strata', and these are then sampled independently. This may be done in small samples to ensure that different groups are represented in proportion to their total share of the population. Alternatively, it may be done so as to deliberately over-sample small groups (e.g. minority ethnic groups) in order to allow meaningful separate analyses of that group. When the individual strata are more homogenous than the total population, conducting analyses within strata may achieve more precise estimates with (legitimately) smaller standard errors [594]. The effect of adjusting for stratification in one's analyses is therefore typically to decrease the standard error slightly.

<u>Clustered sampling</u>

One- and two-phase designs

In many surveys, individuals are not sampled independently but rather are sampled as clusters such as households, schools, or postcodes. Clustered sampling is often logistically a more convenient method of collecting data by, for example, allowing sampling from across a large geographical area at lower cost. The first clusters to be sampled are called primary sampling units. If these are the only clusters selected, the design is called a one-phase design. Alternatively, there may be further sampling of secondary clusters within primary clusters - e.g. households within neighbourhoods. This is called a two-phase design [594].

Clustered sampling may be performed either with stratification (sampling clusters within strata) or without (sampling clusters from the total population). The difference between stratified sampling and clustered sampling is that in stratified sampling the strata are fixed in advance and cover the whole of the population of interest, whereas in clustered sampling the clusters represent a randomly selected sample of groups from the wider population of interest [594].

Design effects from clustering

It is important that if cluster sampling is used then the clustered nature of the data be taken into account during analysis. This is because of the expectation that individuals from the same cluster will be more similar on average than randomly selected individuals in the total population. For example, individuals in the same household may share many environmental exposures, and individuals in the same neighbourhood may have a similar social background.[43] The result is that the variance in a sample obtained through clustered sampling is expected to be lower than the true variance in the population. This will lead to underestimation of the standard error (i.e. √(variance/sample size)), and so to misleadingly narrow confidence intervals, misleadingly large test statistics and misleadingly small p-values. Taking clustering into account is therefore expected to widen the confidence intervals and increase the p-values towards 1. It is not usually expected to alter the point estimates [594-595].

The extent to which failing to adjust for survey design underestimates the variance depends on the 'design effect' [596]. There are two related aspects of the design effect, these being:

1. DEFT: The ratio of the variance of an estimator from the clustered sampling scheme to the variance of the estimator under an assumption of simple random sampling (with the same total number of units).

[43] It is far rarer for individuals in the same cluster to be *less* similar to each other than individuals selected at random from the total population. This does sometimes occur, however, particularly in situations where individuals in a cluster are competing for the same resources. For example, if multiple plants in the same patch of ground are competing for sunlight then taller height in one plant may be correlated with stunted height in another.

2. DEFF: The ratio of the sample size needed for this design compared to the one needed to achieve the same precision under an assumption of simple random sampling.

With DEFT = DEFF^2. For the purposes of analysing data that has already been collected, DEFT is the more relevant measure, and is calculated as follows:

Equation 13.20: Formula for DEFT

$$\text{DEFT} = \sqrt{[1 + (n - 1)\rho]}$$

where *n* is the number of members per cluster and ρ is the intra-class correlation (ICC). As previously in Equation 13.1 (Section 13.1.2), the general formula for the intra-class correlation is:

$$\frac{\text{Between-subject variance}}{\text{(Between-subject variance + within-subject variance)}}.$$

In this case the 'subjects' are the different clusters. The magnitude of the design effect therefore increases as the number of members per cluster increases (which for any given sample size implies a smaller total number of clusters) and/or as the between-cluster variance increases relative to the within-cluster variance.

Sampling with and without replacement

If surveys sample without replacement, then the sample variance tends to increase compared to the variance which would be obtained through sampling with replacement [595]. This is because sampling with replacement means that the same individuals/clusters can be included more than once, which reduces the variance in the sample. This has only a small effect, however, if the sampling fraction of clusters is not large (<10-15%).

Clustered sampling with probabilities proportionate to size

If clusters vary in size, then using simple random sampling to select clusters will lead to small clusters being selected with the same probability as big clusters, despite the fact that they make a smaller contribution to the total population. For this reason, a common strategy is to sample clusters at random with a probability proportionate to their size, and then to sample a fixed number of individuals within each cluster. The advantage of this approach is that if the clusters are sampled with replacement then all individuals in the total population have the same probability taking part in the survey regardless of the size of the cluster they belong to. For example, let *i* be the number of clusters each of size m_h in a total population of size *M*. Let *X* be the fixed number of individuals sampled per cluster. Every time a cluster and its constituent members are sampled, the probability of individual *j* in cluster *i* of being selected is therefore:

Equation 13.21: Effect of sampling clusters with a probability proportionate to their size upon an individual's probability of being represented in the sample.

$$\text{Probability of individual } (i,j) \text{ being selected} = \text{probability of selection of cluster } i \times \text{probability of selection of individual } j \text{ within cluster } i.$$

$$= \frac{m_i}{M} \times \frac{X}{m_i}$$

$$= \frac{X}{M}$$

The probability of selection for all individuals is therefore X/M, a fixed number which does not vary according to cluster size. Note that this requires clusters to be sampled with replacement each time; if this does not occur then individuals from larger clusters will be underrepresented. The degree of this underrepresentation increases as the variation in size between clusters increases and as the sampling fraction of clusters increases.

Probability weights

Sampling clusters of different sizes without replacement is therefore one reason why different individuals in a population can have different probabilities of taking part in a survey. Other reasons include stratification followed by deliberate over- or undersampling, or differential non-response rates between groups. Failing to take account of such differential probabilities of inclusion may lead to biased estimates of prevalence or effect size, if these vary systematically between groups [594-595].

Probability weights (also called sampling weights) attempt to correct for this problem by giving more weight to the results from underrepresented groups. Probability weights are calculated as equal or proportional to the inverse of the probability that an individual from a particular group is sampled. For example, imagine n individuals sampled from a total population of size N. Let $i = 1,\ldots L$ and represent a number of separate groups within the population. Each group contains a total of N_i individuals, and from each group n_i individuals are sampled. The probability weight w_i for individual in group i is the inverse of the probability of individuals from that group being selected. This is given as:

Equation 13.22: Formula for probability weights

$$w_i \propto 1/(n_i/N_i)$$
$$\propto N_i/n_i$$

Conceptually, the value of N_i/n_i therefore corresponds to the total number of individuals in the population represented by each individual from group i. A common practice is to scale N_i/n_i by the total sampling fraction (i.e. n/N), thereby making it more obvious which groups of individuals are over-represented in the sample (weights of <1) and which underrepresented (weights of >1).

Applying weights is expected to alter both the point estimates and the standard errors of means, proportions and effect sizes [594-595]. For example, let $x_1 \ldots x_n$ be a set of observations from individuals independently sampled from a population, each of whom has a probability weight of w_i. The formulae for the weighted mean and variance are:

Equation 13.23: Use of weights in calculating mean and variance

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \qquad V(\bar{x}_w) = \frac{\sigma^2 \sum_{i=1}^{n} w_i^2}{\left(\sum_{i=1}^{n} w_i\right)^2}$$

Where σ^2 is the variance of the population, estimated from the unweighted data in the usual way. Note that this means that when all the weights are equal to 1 (i.e. no weighting), this formula reduces to $\sigma^2 n/n^2 = \sigma^2/n$, which is the basic formula for the standard error. The more the weights are scattered above and below 1, the faster numerator of the equation tends to increase relative to the denominator, leading to an increase in the overall variance. As such, failure to include weighting in one's analysis is like clustering in that it typically leads to an underestimate of the variance of the population.

13.3.2 Adjusting for complex survey design in this thesis

Statistical software

Specialised software for accommodating complex survey design exists in both Stata and MPlus. Both software functions have been explicitly designed to incorporate stratification, clustering and probability weights. Both use methods which adapt the maximum likelihood methods used by standard commands to estimate parameters using a pseudo maximum likelihood (PML) methods [476, 597]. The difference is that for independent, non-weighted data the likelihood to be maximised reflects the joint probability distribution of the data in the chosen mode. If the data are clustered and/or have weights this probabilistic interpretation no longer holds (hence 'pseudo' maximum likelihood). It is, however, still possible to obtain valid parameter estimates after appropriate weighting and after using a method to correct the variance [598]. Because these methods are not genuine maximum likelihood methods, however, they cannot be used to support likelihood ratio tests. The calculation of robust standard errors using the Taylor linearization [598-599] is the default for the 'svy' commands in Stata [476] and is also implemented in MPlus [477].

Application of adjustments for clustered sampling design in this thesis

Throughout this PhD, I allow for the complex B-CAMHS survey design using the complex survey functions in Stata ('svy' family of commands, with the 'strata', 'primary sampling unit' and 'probability weight' options) and MPlus ('Analysis = complex' model specification, with the 'stratification' 'cluster' and 'weight' options). Specifically, I adjust for the complex survey design this whenever calculating proportions and means; when fitting regression models (including those using multiple imputation); when conducting exploratory and confirmatory factor analyses; and when calculating Pearson's correlation coefficients

For two types of analysis, however, adjustment for survey design was not possible. First, neither program currently allows for adjustment for survey design in calculating Spearman's coefficients. Secondly, because the methods used to adjust for complex survey design use robust standard errors and pseudo-maximum likelihood methods, they cannot be used for likelihood ratio tests (see Section 13.2.3).

In the case of likelihood ratio tests, it was therefore not possible to *compare* models while adjusting for survey design. Once the better model had been selected, however, I was able to adjust for survey designing in presenting that model. The stages were therefore as follows:

1. Calculate likelihood ratio of A) nested model and B) general model. ***Not adjusted for survey design.***
2. Use likelihood ratio to determine whether there is any evidence that the nested model is worse. If so retain the general model, otherwise choose the nested model. ***Not adjusted for survey design***.
3. Present results from the chosen model. ***Adjusted for survey design***.

Assessment of effects of adjustments for clustered sampling design

To assess whether these rare instances of non-adjustment for the complex survey design were likely to be a major problem for interpretation of my analyses, I examined the effect of adjusting for the various aspects of the complex B-CAMHS survey design upon the mean parent and teacher SDQ in Whites and Indians. In fact, as shown in Table 13.7, the effect of adjusting for survey design was not large. Compared with the unadjusted error, weighting the data changed the point estimates of the mean only slightly and left the standard errors unchanged. This relatively small effect reflects the fact that the probability weights used in B-CAMHS were small, with 89% lying between 0.085 and 1.15. As expected, adjusting for clustering and stratification left the weighted means unchanged but increased the standard errors and DEFT. This increase was relatively modest, however, (DEFT≤1.35) reflecting low intraclass correlations within clusters (ICC=0.048 for the total sample of parents, 0.020 for teachers).

Based on these findings, I believe that my inability to adjust for complex survey design in a few of the analyses presented in this PhD does not represent an important limitation.

Table 13.7: Effect of adjusting for study design upon point estimates and standard errors for the mean parent and teacher SDQ total difficulty scores (TDS) in Whites and Indians

		N	Point estimate	Standard error	DEFT
Mean parent TDS score in Whites	**Unadjusted**	16402	8.249	0.0458	1
	Plus weighting	16402	8.247	0.0461	1.01
	Plus clustering and stratification	16402	8.247	0.0604	1.32
	Plus finite-population correction	16402	8.247	0.0587	1.28
Mean parent TDS score in Indians	**Unadjusted**	390	7.594	0.265	1
	Plus weighting	390	7.614	0.277	1.02
	Plus clustering and stratification	390	7.614	0.341	1.30
	Plus finite-population correction	390	7.614	0.331	1.26
Mean teacher TDS score in Whites	**Unadjusted**	12913	6.506	0.0528	1
	Plus weighting	12913	6.501	0.0532	1.01
	Plus clustering and stratification	12913	6.501	0.0579	1.10
	Plus finite-population correction	12913	6.501	0.0562	1.06
Mean teacher TDS score in Indians	**Unadjusted**	306	5.229	0.273	1
	Plus weighting	306	5.216	0.275	1.02
	Plus clustering and stratification	306	5.216	0.348	1.31
	Plus finite-population correction	306	5.216	0.338	1.27

Options not used: Finite population corrections

The default assumption in adjusting for survey design in both Stata and MPlus is that primary sampling units were sampled with replacement. This was not the case in the B-CAMHS surveys, and I therefore applied the finite-population correction option of the svy family. This typically specifies the total number of primary sampling units (i.e. postal sectors) per stratum in the population. Unfortunately, ONS were not able to provide me with the precise number of postal sectors in each of the 231 strata in B-CAMHS99 and B-CAMHS04. Instead, I could only estimated the average numbers in each country, as shown in Table 13.8. The estimated number of sectors per strata in columns D and F were used as the finite-population correction factors for all strata in a particular country in each year.

Table 13.8: Estimation of number of sectors per stratum in B-CAMHS99 and 04

	A	B	C	D	E	F
	% postal sectors in country	**Estimated total no. postal sectors (A*8265/100)†**	**No. strata in B-CAMHS99**	**Est. average sectors/ stratum B-CAMHS99 (B/C)**	**No. strata in B-CAMHS04**	**Est. average sectors / stratum B-CAMHS04 (B/E)**
England	86.4	7140.96	196	36	185	39
Scotland	8.72	720.71	24	30	18	40
Wales	4.95	409.12	12	34	5	82
Total	100	8265	231	36	208	40

†Source ONS documentation for B-CAMHS99 [2, Appendix A] This gives the total number of postal sectors in the UK (N=8265) plus the percentage of sectors in each country of Great Britain.

In Stata it is possible to use these estimates to make finite-population corrections to the standard error. This facility has not yet been added to MPlus (Linda Muthen, personal communication). In fact, applying these finite-population corrections reduced the standard error only very slightly (<3%), as shown in Table 13.9. This reflects the large estimated number of clusters per stratum (N=34-82; see Table 13.8) with only two or three clusters were sampled from each stratum. Moreover, as expected, the effect was to decrease the standard error, meaning that my failure to adjust for sampling without replacement will lead to conservative estimates of test statistics.

Thus finite population corrections in Stata were based only upon estimates of sectors per strata and were not available for analyses in MPlus. Moreover, failure to make finite population corrections had only a small effect in a conservative direction. For all these reasons I decided not to use finite population in this thesis.

Table 13.9: Effect of additionally adjusting for sampling with replacement upon point estimates and standard errors for the mean parent and teacher SDQ total difficulty scores in Whites and Indians

		N	**Point estimate**	**Standard error**	**Percent change in standard error**
Mean parent total difficulty SDQ score in Whites	**Adjusted for weighting, clustering and stratification**	16402	8.247	0.0604	
	Plus finite-population correction	16402	8.247	0.0587	2.9% decrease
Mean parent total difficulty SDQ score in Indians	**Adjusted for weighting, clustering and stratification**	390	7.614	0.341	
	Plus finite-population correction	390	7.614	0.331	2.9% decrease
Mean teacher total difficulty SDQ score in Whites	**Adjusted for weighting, clustering and stratification**	12913	6.501	0.0579	
	Plus finite-population correction	12913	6.501	0.0562	2.1% decrease
Mean teacher total difficulty SDQ score in Indians	**Adjusted for weighting, clustering and stratification**	306	5.216	0.348	
	Plus finite-population correction	306	5.216	0.338	2.8% decrease

‘Single’ clusters within a stratum

Stata 9.2 requires at least two clusters per stratum in order to calculate the stratum-specific between-cluster variance. In theory, all clusters in the B-CAMHS dataset were paired to one (sometimes) two other clusters in a stratum. In a minority of subset analyses, however, all individuals from one or more clusters were outside the population or interest or had missing data. The clusters to which they were paired therefore became ‘single’ for these analyses. In these cases, I excluded the individuals in these single clusters. The fact that this only ever applied to a tiny minority of individuals (0-0.6%) means this is not expected to have had any effect on my substantive findings.

13.4 Multiple Imputation for missing covariate data

13.4.1 Types of missing data

Data for particular items can be missing for participants in a survey for many reasons, including incomplete responses, data entry errors, loss to follow-up or an item not being included in a particular round of a survey. In all cases this leads to loss of power, and in some circumstances it may lead to biased estimates of effect. To assess whether this is likely to be an issue, and what the best strategy is for dealing with missing data, it is important to understand the missingness mechanism – that is, the pattern of missingness of the data.

To this end, one can distinguish three broad classes of missingness mechanisms [600-601]. Each can be described as a joint function of the values of the observed data (Y_{OBS}); the values of the missing data items (Y_{MISS}); and a 'missingness indicator' R which takes the value 1 if the item is missing and 0 if it is not.

1. **Missing Completely At Random (MCAR):** The probability that a data item is missing is not a function either of the data you have observed or of the data you are missing. That is, the missingness mechanism is as follows:

   $$[R \mid Y_{OBS}, Y_{MISS}] = [R]$$

   In the case of MCAR, estimates using only individuals in the dataset with complete information will give an unbiased estimate (although this may still be inefficient due to the reduction in sample size caused by eliminating individuals with near-complete data).

2. **Missing At Random (MAR):** The probability that a data item is missing is a function of the data you have observed but, after conditioning on the observed data,

is not a function of the data you are missing. That is, data is selectively missing, but this is for reasons entirely explained by other variables you have observed, as follows:

$$[R \mid Y_{OBS}, Y_{MISS}] = [R \mid Y_{OBS}]$$

In describing instances of MAR, it is therefore important to specify which variables are included within Y_{OBS}; for example 'missing at random *given gender and age*'. If MAR is the missingness mechanism, then listwise deletion of individuals may give biased estimates, but this can be corrected through model-based methods of imputation.

3. **Not Missing At Random (NMAR):** The probability that a data item is missing is a function of data which you are missing. That is:

   $$[R \mid Y_{OBS}, Y_{MISS}] \text{ cannot be simplified}$$

   If NMAR is the missingness mechanism, then listwise deletion of individuals may give biased estimates, and this cannot be corrected using model-based approaches.

Clearly in real life many missingness mechanisms may be NMAR. It is, however, still useful to use model based corrections based on an assumption of MAR as these will tend to give less biased answers than ignoring the missing data completely. Furthermore, analyses based on MAR act as a useful point of departure for any subsequent NMAR sensitivity analyses.

13.4.2 Multiple imputation as a strategy for analysing missing data

Rationale for multiple imputation

In analysing data, it is preferable to use a principled approach using statistical models in which each person makes an appropriate contribution based on their observed data. In the case of missing units (e.g. people not measured at all) one can use probability weights to weight back ones results to the original sample, as discussed in Section 13.3.1. In the case of missing data items, multiple imputation is the preferred strategy [534], particularly in instances in which data is missing for covariates as well as outcomes [602]. Multiple imputation involves fitting a model in which the value of each missing data item is predicted (imputed) based upon the observed data, under an assumption of MAR. This can be done using draws from a model for the variable in question which uses the values of observed variables to impute the missing values. In performing such an imputation, one is not creating 'plausible' values but rather drawing from the underlying distribution of a variable in terms of both its mean and variance.

One then combines results from across multiple such imputations in order to fit a model of substantive interest. Combining results from across multiple imputations is essential because in imputed data there are two sources of variation; the variation in the underlying variable of interest and the variation (uncertainty) in the imputation process. A single imputation model therefore gives unbiased estimators but, by treating imputed values as if they were in fact observed, it ignores the second source of variation. It therefore generates standard errors which are too small. Combining results from across multiple imputations provides a means of modelling the additional uncertainty associated with the imputation process and so correcting these standard errors. Five imputation models has typically been regarded as adequate for most purposes, although this may increase if one wants to go beyond Wald tests of single comparisons (i.e. one level vs. baseline) and instead perform a joint test of all categories in a variable (e.g. 'ethnicity' containing multiple ethnic groups) [603].

Multiple imputation by chained equations (MICE)

In this thesis I conduct multiple imputation by chained equations [533-534], using the MICE (multiple imputation by chained equations) command in Stata10 [535-536], and subsequently combining and analysing the datasets using the MIM command [604]. MICE uses regression analyses to impute each missing value in the dataset based on all other variables in the model. Each variable with any missing data is taken in turn as an outcome, and the observed data is used to fit a regression model for that variable based on all the observed data. Under an assumption of missing at random, this relationship is assumed to be the same for the missing values. The model therefore provides an estimated regression line for the values of the missing variables, plus estimated sampling distributions for the values of the regression coefficients and the standard deviation of the outcome.

Imputed values cannot simply be drawn at random from around this line for the missing values of the current variable of interest. This would represent an improper imputation as it would only capture the sample variance and not the extra variance created through the fact of imputing (rather than observing) data. For this reason the imputed values are instead drawn at random from around a line which is itself based on random draws from the Bayesian posterior distributions of the parameters that defined the fitted regression line. The result is that Rubin's simple variance formula for estimating the standard errors then gives a consistent estimate of the true standard error of the parameter estimates from the multiple imputation analysis [534].

The imputation model then proceeds to the next covariate of interest, and this time uses imputed values from the first covariate as part of the estimation process. In this way, the model cycles through all the variables in turn until all missing values are filled in, this being the 'chained equation' (and also known as a Gibbs sampling type method).

Content of the imputation model

The multiple imputation model should contain two broad types of variable:

1. **All components of subsequent substantive models of interest**, including all outcome variables, explanatory variables and all model structure (e.g. quadratic terms or potential interactions). This is important as if one does not include model structure such as interactions in the imputation model then one will be less able to detect such an effect in the model, as the imputation would be performed based on a model assuming no such interaction exists. Likewise if you do not include a covariate from the substantive model of interest in the imputation structure, you may bias the estimated effect towards the null [534].
2. **Anything that is part of missingness mechanism**; that is, anything which predicts which individuals have missing data *or* which predicts the missing values of a particular covariate. Because the imputation model needs to include all components of the substantive model of interest and also all components of the missingness

mechanism, it will usually be larger than the substantive model of interest. Note that the missingness mechanism may vary between covariates, and that this does not matter.

In general, the more covariates which one includes in the imputation model the closer one will come to meeting the 'missing at random' assumption. As such, an inclusive strategy is generally preferable to a restrictive strategy when fitting imputation models – particularly given that, unlike standard 'model building' there is relatively little advantage to being parsimonious.

Chapter 14 Appendix 2: Supplementary information on analyses presented in thesis

14.1 Appendix to Chapter 4: Supplementary information on systematic review

Supplementary information on search methods

Box 14.1: Electronic search methods

Search string
(#Infant OR #Child OR #Adolescent) AND (#Mental health OR #Mental disorder OR #Psychiatry OR #Psychiatric clinics OR "mental health" OR "mental illness" OR "mental distress" OR "mental disorder*" OR "behaviour disorder*" OR "behavior disorder*" OR "emotional disorder*" OR "hyperactiv*" OR "hyperkinesis" OR externali*ing OR internali*ing OR anorexia OR bulimia OR "eating disorder*" OR self-harm OR self-injur* OR suicid* OR somatoform OR autism OR autistic OR psychosis OR psychotic OR psychoses) AND (#Britain OR #United Kingdom OR "British" OR "Britain") AND (#Ethnic groups OR #Ethnic differences OR #Minority groups OR#Cross-cultural comparison OR Ethnic* OR Migrant* OR Immigrant* OR Minorit* OR Race OR Racial OR "cross-cultural" OR "cross cultural" OR White OR Caucasian OR mixed race OR Black OR African OR Afro OR West Indian OR Asian OR Indian OR Bangladeshi OR Bengali OR Pakistani OR Punjabi OR Gujarati OR Tamil OR Chinese)

#=indicates exploded indexing terms to include all subheadings. *=wildcard symbol. Search terms in quotes searched as exact phrases.

Databases (and search date)
BNI* (25/5/07, updated 14/7/07)
British Library theses index† (25/5/07, updated 19/7/07)
CAB Direct (22/5/07, updated 11/7/07)
CINAHL (23/5/07, updated 14/7/07)
Cochrane (22/5/07, updated 11/7/07)
DH National Research Register* (23/5/07, updated 14/7/07)
DH ReFeR* (22/5/07, updated 14/7/07)
Embase (6/5/07, updated 11/7/07)
ESTAR† (25/5/07, updated 19/7/07)
HMIC (22/5/07, updated 12/7/07)
IBSS (22/5/07, updated 11/7/07)
PubMed (23/1/07 and 17/2/07[44], updated 11/7/07)
PsycInfo (16/2/07, updated 11/7/07)
Science and Social Science Citation Index (22/5/07, updated 11/7/07)
TRIP† (25/5/07, updated 12/7/07)
Zetoc† (25/5/07, updated 19/7/07)

*='Britain'-related terms omitted from search strings. †=key word searching used as the database only supported a limited number of search terms.

Websites
Centre for Evidence in Ethnicity, Health and Diversity (CEEHD) [605]
Centre for Research in Ethnic Relations [CRER] [606].

[44] Searched twice to assess reliability - see text.

The NHS Specialist Library for Ethnicity and Health [607].
Confederation of Indian organisations [608].
The Department of Health National electronic library for health [609].
The King's Fund [610].
Mind (National Association for Mental Health) [611].
The National Ethnic Minority Data Archive (NEMDA) [612].

Special interest groups and e-mail distribution lists
The Royal College of Psychiatry's Transcultural Psychiatry special Interest Group.
The Royal College of Psychiatry's faculty of Child and Adolescent Psychiatry.
CAMHS@JISCMAIL.AC.UK
YOUNGPERSONS-PSYCHIATRIC-NURSING@WWW.JISCMAIL.AC.UK
MINORITY-ETHNIC-HEALTH@JISCMAIL.AC.UK

Supplementary information on inclusion criteria

'Comparable' general population

The minimum criteria for a general population sample to be considered 'comparable' were: matched for population vs. clinic-based; nationally representative sampling or representative sampling from the same geographic area (operationalised as a Government Office Region); same mental health outcome(s) using (where applicable) the same interviews/ questionnaires and cut-offs; and matched for year of data collection to ±10 years.

Minimum sample sizes

The minimum sample sizes for the three survey designs included in this review were determined as follows.

- Population-based prevalence or mean score studies: minimum group size N≥40 for prevalence for each included ethnic group, corresponding to a difference of 10 vs. 25% (for questionnaire cut-offs) or 10% vs. 3% (for disorder prevalences). N≥10 for mean scores for each included ethnic group, based on effect size of 0.7.
- Clinic-based studies of the relative proportion of referrals/in-patients in clinics from ethnic minority groups: No minimum group size.
- Clinic-based studies of the proportional morbidity of different disorders: minimum group size N≥20 for each included ethnic group, based on moving from a 50:50 split of e.g. internalising and externalising disorders to a 75:25 split.

These minimum ethnic group sizes are based on power calculations with significance at 5% and power set at 50% – i.e. the level at which half of genuine differences will be missed. This, in combination with the assumption of relatively large ethnic differences, deliberately sets the bar for inclusion quite low. The power calculations make allowance for the possibility that a minority group may only make up a small proportion of the total sample population.

Supplementary information on search results

Table 14.1: Electronic sources searched and hits returned

Electronic sources	**No. hits**	**No. new hits of some relevance**	**No. included in the review as a primary or secondary study**
DATABASES (listed in order of searching)			
PubMed	1391	30 [376, 379-381, 383-384, 389, 411, 415, 418, 556, 613-631]	11 [376, 379-381, 383-384, 389, 411, 415, 418, 628]
PsycINFO	1316	15 [369, 382, 385-387, 390, 392, 397, 409, 447, 541, 632-635]	8 [382, 385-387, 390, 392, 397, 409]
Embase	556	1 [636]	0
CAB Direct	108	0	0
Cochrane	6	0	0
DH ReFeR	32	0	0
HMIC	69	0	0
IBSS	31	0	0
Science and social science citation index	404	1 [637]	0
CINAHL	104	0	0
DH National research register	1117	8 [638-645]	0
British Library theses index	3	0	0
BNI	85	0	0
ESTAR	59	0	0
TRIP	30	0	0
Zetoc	33	0	0
CITATIONS (combined)	808	4 [646-649]	0
WEBSITES			
Centre for Evidence in Ethnicity, Health and Diversity [605]	n/a[1]	0	0
Centre for Research in Ethnic Relations [606]	102	0	0
NHS Specialist Library for Ethnicity and Health [607]	5	0	0
Confederation of Indian organisations [608]	n/a	0	0
DH National electronic library for health [609]	23	0	0
King's Fund [610], including their 'ethnic health' reading list [650].	4	0	0
Mind (National Association for Mental Health) [611]	n/a	0	0
National Ethnic Minority Data Archive [612].	0	0	0
TOTAL	**6286**	**59**	**19**

1'N/a' refers to websites which were browsed not searched, as no formal search functions existed.

Box 14.2: Methodological limitations

Measure of mental health

- A: Mental health information only from a questionnaire measure.
- B: Reliance on a single, inappropriate informant. Inappropriate informants were defined as a) self-report by child aged under 11 [651-652], b) self-report where oppositionality and hyperactivity are the key outcomes (anti-social behaviour scales accepted) [653], or c) teacher-report where emotional disorders, eating disorders or deliberate self-harm are the key outcomes [2, Appendix A].
- C: Non-validated modification of a validated mental health score.

Measurement of ethnicity

- D: Method of assigning ethnicity not described.
- E: Ethnicity determined by a potentially inferior method. Adequate methods were child's self report, parent's report on child, or according to parent's/grandparent's country of birth/'country of origin. Other methods were considered inferior – for example using child's place of birth, ethnicity as assigned by name, or ethnicity as ascribed by clinicians. Where case notes were used to assign ethnicity, but it was not explicitly stated how information on case notes was completed, this was recorded ?E.
- F: Ethnicity analysed using meta-level descriptions only ('White' (combining White British/Irish with White minority), 'Mixed', 'Black' or 'Asian'/'South Asian') or by comparing one ethnic group to a potentially mixed-ethnicity comparison group (e.g. all other children in the sample/the 'general population').

Methodological limitations of the study that may cause bias

- G: Potential for selection bias – clinic-based sampling.
- H: Potential for selection bias – response rates less than 60% for population-based surveys or completeness of ethnicity data less than 60% for clinic surveys.
- I: Potential for information bias – investigator-based ratings made by the study authors without being blinded to ethnicity.

Alternative explanations for observed differences

- J: Differences could be due to confounding by age and sex (including cases where no information was given). This was taken to apply unless a) study was restricted to one sex or age range ≤ 3 years b) similar age (< 1 year difference) and/or sex profile (<10% difference) was demonstrated in the different ethnic groups, c) results were stratified by age (age bands ≤ 3 years) and/or sex, or d) age and/or sex were controlled for in multivariate analyses. This criterion was NOT applied to clinic based referral rate studies – i.e. it was assumed that the base population was balanced in its age and sex composition.
- K: No data presented on socio-economic position
- L: No adjustment made for reported differences in socio-economic position.

Table 14.2: Studies not included in review, with reasons for non-retrieval or non-inclusion

	Author, date	Details
Unable locate or research described never carried out	Arnone [644]	Apparently unpublished research proposal, author(s) could not be contacted.
	Minnis in press [654]	Unpublished study which had been cited [in 362] but where the authors were not clear what study was meant [655].
	Ramjee [645]	Apparently unpublished research proposal, author(s) could not be contacted.
	Skinner [642]	Apparently unpublished research proposal, author(s) could not be contacted.
	Walker 1968 [656]	Unpublished, could not be located.
	Zaineb [638]	Apparently unpublished research proposal, author(s) could not be contacted.
	Zietlin [640]	Apparently unpublished research proposal, author(s) could not be contacted.
Data collection/analysis not completed before deadline	Hodes [641]	Data collection or analysis not completed before the deadline.
	Kelly [657]	Data collection or analysis not completed before the deadline.
	Laurens [658]	Data collection or analysis not completed before the deadline.
	Muthulagu [643]	Data collection or analysis not completed before the deadline.
	Truman [659]	Data collection or analysis not completed before the deadline.
Sampled from specially selected groups	Chowdhury 2005 [615]	Sampled from a medium secure health service unit, mostly serving adolescents convicted of a criminal offence.
	Kelsall 1995 [660]	Sampled from a medium secure health service unit, mostly serving adolescents convicted of a criminal offence.
	Leavy 2004 [630]	Compares refugee to non-refugee children.
	Morita 1987 [623]	Children of expatriate Japanese businessmen, temporarily studying in Japanese schools in Britain.
	Ullah 1985 [636]	Unemployed 17-year olds receiving benefits.
Did not contain an eligible ethnic group	Ahmad 1994b[646]	South Asian population subdivided by religious group (Hindu and Muslim). Note: results on female sample are presented in Ahmad 1994a
	Boeing 2007 [629]	Only ethnic comparison is European vs. non-European origin.
	Burt 2004 [614]	Only ethnic comparison is White British vs. non-White British.
	Flouri 2006 [618]	Only ethnic comparison is White vs. non-White.
	Gowers 1993 [619]	Only ethnic comparison is White British vs. non-White British.
	Mears 2003 [635]	Only ethnic comparison is White vs. non-White.
	O'Herlihy 2001 [661]	Only ethnic comparison is White vs. non-White.
	Schmidt 1992 [662]	Only 'ethnic' comparison is according to membership of an ethnically heterogeneous group loosely defined by 'cultural stress'
	Sonuga-Barke 2000 [539]	Compares Hindu children with Muslim children, this division cutting across families India, Pakistan or East Africa as a place of origin.
	Stevenson 1985 [625]	Only ethnic comparison is immigrant vs. non-immigrant.
Contained no eligible ethnic comparison	Lindsey 2003 [634]	Inadequate information about the small external general population comparison group to determine when and where it was collected.
	Nicol 1971 [1st study] [663]	Comparison within Black-Caribbeans, according to whether they were born inside the UK or not.
	Shah 1995 [624]	Comparison within Pakistanis, according to whether they were from extended or nuclear families.
Did not give the number	Dex 2007 [540]	Number of children of each ethnic group not given
	Cochrane 1979	Number of children of each ethnic group not given

of participants per ethnic group	[2nd Study] [373]	
	Furnham 1994 [664]	Inconsistencies between number of children as reported by ethnic group and as reported by religion, which the authors could not clarify [665].
	Hillier 1994 [666]	Brief reference to an underrepresentation of Bangladeshis at their child and adolescent mental health clinic in passing.
	Vyas 1998 [639]	Brief poster abstract in which number of children of each ethnic group not given.
Did not measure a mental health outcome of interest	Moran 2007 [667]	Mental health outcome outside the scope of this review (psychopathy).
	Harris 2001 [668]	Evaluation of service delivery
	Hill 1995 [621]	Measures eating style/behaviour, but not eating disorders.
	Ogden 1998 [648]	Measures eating style/behaviour, but not eating disorders.
Did not use a valid/ validated mental health measure	Bhatnagar 1970 [669]	Non-validated 'adjustment scale' constructed for the study.
	Dosanjh 1976 [670]	Non-validated mental health and temperament scales constructed for the study.
	Dosanjh 1996 [555]	Non-validated mental health and temperament scales constructed for the study.
	Hackett 1993 [556]	Non-validated question on 'temper tantrums', constructed for the study.
	Schools Council 1970 [671]	Measures social adjustment using the Bristol Social Adjustment Guide, [672] judged not to be satisfactorily validated as a mental health measure.
	Hawton 2002 [673]	Did not use a validated instrument for assessing deliberate self harm
	Marchant 2006 [637]	Mental health status assessed using school registries of special educational needs for autistic spectrum disorder – i.e. an administrative registry not created for the purposes of mental health research.
Not of an eligible study type or size	Bhugra 2002 [647]	Compares reasons and risk factors among those who have committed deliberate self-harm.
	Earls 1980b [616]	Compares mental health prognosis.
	Handy 1991 [620]	Compares methods of self-poisoning and risk factors among those admitted with deliberate self-harm.
	Holden 1988 [622]	Compares symptoms within matched samples of children with anorexia.
	Kingsbury 1994 [649]	Compares symptoms and risk factors within children who have committed deliberate self-harm.
	Nicol 1971 [2nd study] [663]	Compares mental health prognosis.
	Nicol 1971 [3rd study] [663]	Compares symptoms within children with behavioural disorder
	Sonuga-Barke 1993 [1st study] [232]	Compares teacher's assessment of hyperactivity against objective measures.
	Sonuga-Barke 1993 [2nd study] [232]	Compares teacher's assessment of hyperactivity against objective measures.
	Tareen 2005 [626]	Comparing risk factors and symptoms within girls presenting with eating disorder.
	Vostanis 2003 [627]	Compares service use among children with conduct disorder
Referral rates not calculated in relation to a specified base	Bendall 1972 [633]	No explicit comparison of number of diagnoses of anorexia in South Asians with a base population.
	Fatimilehin 1998 [674]	No explicit comparison of number of referrals for Black and non-Black children with a base population.
	Subotsky 1990	No explicit comparison of number of referrals from each ethnic

population	[675]	group with a base population.
Could not extract an analysis by ethnic group for an outcome of interest	Best 2006 [631]	Does not present mental health outcome by ethnic group.
	Daryanani 2001 [447]	Single simultaneous comparison of the proportional morbidity for 19 mental health outcomes, only 7 of which are within the scope of this review and which cannot be extracted from those which are not.
	Shams 1995 [676]	Mental health outcome of interest (GHQ) not compared between ethnic groups, but used only as a covariate.
Duplicate publications containing no new information.	Atzaba-Poira 2004b [632]	Duplicates Atzaba-Poira 2004a [387]
	Bhui 2005 [613]	Duplicates Stansfield 2004 [388]
	Bhugra 2004 [369]	Duplicates Bhugra 2003b [408]
	Deater-Deckard 2004 [544]	Duplicates Atzaba-Poira 2004a [387]
	Earls 1982 [617]	Duplicates Earls 1980a [376]
	Hackett 1994 [541]	Duplicates Hackett 1991 [377]
	Mumford 1988 [677]	Duplicates Mumford 1991 [378]

Table 14.3: Description of studies[45]

	All studies			Studies with Indian samples		
	Population-based	**Clinic-based**	**Total**	**Population-based**	**Clinic-based**	**Total**
Decade of study publication						
1960s	0	1	1	0	0	0
1970s	4	0	4	2	0	2
1980s	2	2	4	0	0	0
1990s	9	6	15	2	0	2
2000-2007	16	9	25	9	1	10
Age of study population						
0-10 age range	9	0	9	2	0	2
11-19 age range	12	3	15	5	0	5
Spans both age ranges	10	15	25	6	1	7
Study setting						
Great Britain (nationally representative)	3	0	3	2	0	2
London	15	11	26	4	0	4
England, not London	13	6	19	7	1	8
Scotland	0	1	1	0	0	0
Wales	0	0	0	0	0	0
Mental health outcome						
Any diagnosis/clinic referral for any problem	3	12	15	3	1	4
Common mental health problem	21	6	27	10	0	10
Eating disorders	7	0	7	1	0	1
Psychosis	1	5	6	0	0	0
Deliberate self harm	0	7	7	0	0	0
Other disorders	0	3	3	0	0	0
Ethnic group						
White British	16	3	19	8	0	8
‘White’	12	8	20	3	1	4
White minority	5	0	5	4	0	4
Mixed race	6	0	6	3	0	3
Black Caribbean	13	4	17	6	0	6
Black African	7	1	8	5	0	5
‘Black’ as a meta-level ethnic group	5	6	11	1	1	2
Indian	13	1	14	13	1	14
Pakistani	7	2	9	7	1	8
Bangladeshi	7	4	11	6	1	7
‘South Asian’ as a meta-level ethnic group	12	9	21	0	0	0

[45] Only the 49 independent studies are listed here, to avoid double counting. Scores in some categories total more than 49 because studies often fitted into several categories - for example, containing samples from several ethnic groups.

	All studies			Studies with Indian samples		
	Population-based	**Clinic-based**	**Total**	**Population-based**	**Clinic-based**	**Total**
Chinese	2	0	2	2	0	2
Orthodox Jew	1	0	1	0	0	0
Study limitations						
A: Questionnaire measure of mental health only.	22	0	22	10	0	10
B: Reliance on a single, inappropriate informant.	1	0	1	0	0	0
C: Non-validated modification of a measure.	4	0	4	1	0	1
D: Method of assigning ethnicity not described.	7	1	8	2	0	2
E: Ethnicity determined by an inadequate means.	2	14	16	0	1	1
F: Ethnicity analysed using a meta-level grouping	21	14	35	6	1	7
G: Potential for selection bias – clinic-based sampling.	0	18	18	0	1	1
H: Potential for selection bias – response rates/data completeness <60%.	12	10	22	3	1	4
I: Potential for information bias – investigator-based ratings not blinded to ethnicity.	6	0	6	0	0	0
J: Possibility of confounding by age and sex.	8	2	10	3	0	3
K: No data presented on SEP.	13	15	28	3	1	4
L: No adjustment for reported differences in SEP.	8	2	10	2	0	2
Total number of independent studies	31	18	49	13	1	14

Table 14.4: Description of population-based studies of child mental health

Study ref.	Pub. type	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Bagley 1972 [374]	1	London, ??	Cross-sectional: 186 children aged 7: 112 White British/Irish; 74 Black Caribbean	Not specified	Common mental health problems: teacher	Black Caribbean ↑	A, D, F, ?H, ?J (sex), K
Rutter 1974 [238]; 1975 [415]	1; 1	London, 1970	2-phase cross-sectional: 2 043 children aged 10: First phase: 1689 White; 354 Black Caribbean. Second phase: 265 White; 118 Black Caribbean	Country of origin	Common mental health problems/disorders: teacher and (second phase) parent	Black Caribbean ↑/—	F, L
Kallarackal 1976 [375]	2	Leicester, ??	Cross-sectional: 198 children aged 9-12: 98 White British; 100 Indian	Not specified	Common mental health problems: parent and teacher	Indian ↓*	A, D, L
Cochrane 1979 [373]	1	Birmingham, 1976	Cross-sectional: 301 children aged 9: 74 White British; 87 Black Caribbean; 42 Pakistani; 98 Indian	Country of origin	Common mental health problems: teacher	Black Caribbean ?↑; Pakistani ?↓; Indian ?↓	A
Earls 1980 [376]	1	London, 1972-73	Cross-sectional: N=763 children aged 3: 705 White British; 58 Black Caribbean	Country of origin	Common mental health problems: parent	Black Caribbean —*	A, I, L
Osborn 1985 [372]	2	Great Britain, 1975	Prospective cohort study 12 335 children aged 5: 11907 White British; 187 Black (>90% Black Caribbean); 241 South Asian	Country of origin	Common mental health problems: parent	Black ?↑; South Asian ?↑/—	A, F, H (for black children)
Hackett 1991 [377]	1	Manchester, ??	Cross-sectional: 200 children aged 4-7: 100 White British;100 Gujarati Indian	Not specified	Common mental health problems/referral: parent	Indian ↓*	D, H, ?J (sex), L
Mumford 1991 [378]	1	Bradford, ??	2-phase Cross-sectional: 559 girls aged 14-15: 255 White; 204 South Asian	Not specified	Problematic eating attitudes/eating disorder: self-report and clinical diagnosis	South Asian ↑	D, F, I, K
Newth 1993 [379]	1	Birmingham, 1987	Cross-sectional: 194 children aged 3-4: 65 White; 129 South Asian (111 Pakistanis; 18 Indians)	Not specified	Behavioural problems/disorders: parent and clinical rating	South Asian ↓/—*	?E, F, I, ?J (sex), L
Ahmad 1994 [380]	1	Bolton, 1993†	Cross-sectional: 186 girls aged 14-15: 115 White British; 71 South Asian	Reported by child	Problematic eating attitudes: self-report	South Asian —	A, F, K

Study ref.	Pub. type	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
McCourt 1995 [381]	1	Birmingham, ??	Cross-sectional: 336 girls aged 12-16: 158 White; 178 South Asian	Not specified	Problematic eating attitudes: self-report	South Asian ↑	A, D, F, K
Waller 1995 [382]	1	Birmingham, ??	Cross-sectional: 260 girls aged 14-15: 107 White; 153 South Asian	Not specified	Problematic eating attitudes: self-report	South Asian —	A, C, D, F, K
Marks 1997 [383]	1	London, ??	Cross-sectional: 174 children aged 5-11: 61 White British; 113 Bangladeshi	Not specified	Common mental health problems/disorders: teacher and clinical rating	Bangladeshi —*	B, D, ?H ?I, K
Nikapota 1998 [214]	3	London, 1996-97	Cross-sectional: 258 children aged 9-12: 60 White British; 40 Mixed White-Black Caribbean; 60 Black Caribbean;[38 Black African]; 60 South Asian	Reported by child's parent	Common mental health problems/disorders: parent and teacher	Mixed race ?—; Black Caribbean ?↑/—; South Asian ?↓/—	F, ?H, ?I, J (sex), L
Nazroo 1999 [343]	2	England, 1999	Cross-sectional: 1914 children aged 4-15 : 342 Irish; 363 Black Caribbean; 296 Indian; 412 Pakistani; 319 Bangladeshi; 182 Chinese. Compared with a 1997 general population sample of 5705 children.	Reported by child/child's parent	1) Common mental health problems: parent. 2) emotional problems (for those aged 13-15): self-report	White Irish ↑/—; Black Caribbean —; Indian ↑/—; Pakistani ↑/—; Bangladeshi —; Chinese —.	A, F, J (age), K
Meltzer 2000 [2]; Ford 2003 [416]; Ford 2004 [417]; Evans 2004 [418]	2; 1; 1; 1	Great Britain, 1999	Cross-sectional: 10 438 children aged 5-15: 9529 White; 105 Black-Caribbean; 74 Black African; 222 Indian; 147 Pakistani; 43 Bangladeshi; [34 Chinese; 284 Other]	Reported by child's parent	Common mental disorders: clinical rating.	Black Caribbean—**; Black African —**; Indian ↓**; Pakistani —**; Bangladeshi —**	F
Furnham 2001 [384]	1	London, 1998	Cross-sectional: 168 girls aged 15-17: 46 White; 40 Indian; 44 Pakistani; 38 Bangladeshi	Country of origin	Problematic eating attitudes: self-report	South Asian ↑ (Indian ?—; Pakistani ?—; Bangladeshi ?↑)	A, ?F, K
Thomas 2002 [385]	1	Bristol, ??	Cross-sectional: 653 children aged 11-16: 405 White; 19 Mixed race; 101 Black; 103 South Asian; [25 Other]	Reported by child	Problematic eating attitudes: self-report	Mixed race ↑; Black —; South Asian ↑	A, F, ?H, J (age), K
Bhugra, Bhui 2003 [386]	1	London, ??	Cross-sectional: 266 children aged 13 or more: 134 White British; 51 Black Caribbean; 52 South Asian; [29 Other]	Reported by child	Problematic eating attitudes: self-report	Black —; South Asian —	A, C, F, ?H, K

Study ref.	Pub. type	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Atzaba-Poira 2004 [387]; 2005 [420]; 2007 [421]	1; 1; 3	London, 2001-02; 2005-06	Repeat cross-sectional: Time 1: 125 children aged 7-9: 59 White British; 66 Indian. Time 2: 68 children aged 9-13 (58 from Time 1): 37 White British; 31 Indian	Country of origin	Common mental health problems: parent and, in time 1 only, teacher	Time 1: Indian —/↑ Time 2: Indian —/↑	A, ?H
Stansfeld 2004 [388]; Klineberg 2006 [419]; Fagg 2006 [157]	1; 1; 1	London, 2001	Cross-sectional: 2790 children aged 11-14: 581 White British; 161 White minority; 194 Mixed race; 575 Black; 250 Indian; 184 Pakistani; 690 Bangladeshi; [115 Other]	Reported by child	1) Common mental health problems and 2) emotional problems: self-report	White minority — /↑; Mixed race —; Black —; Indian —; Pakistani —; Bangladeshi ↓/—	A, F (for Fagg 2004)
Edmunds 2005 [389]	1	London, 2003	Cross-sectional: 163 children aged 5: 78 White; 18 Mixed; 34 Black; 33 South Asian	Reported by child's parent	Common mental health problems: parent	Mixed —***; Black —***; South Asian —***	A, F, H
Flouri 2005 [391]	1	South England, 2001-02	Cross-sectional: 582 children aged 11-19: 360 White British; 222 Indian	Reported by child	Common mental health problems: self-report	Indian ↓	A
Frosh 2005 [390]	1	London, ??	Cross-sectional: 341 children aged 5-15: Parent report: 161 Orthodox Jews; 10 298 children from a nationally representative general population sample collected in 1999. Teacher report: 325 Orthodox Jews, 8 028 general population.	Other method	Common mental health problems: parent and teacher	Orthodox Jew ↓/—*	A, E, F, H (for parents), J (age), L
Green 2005 [3]	2	Great Britain, 2004	Cross-sectional: 7 974 children aged 5-16: 6787 White British; 134 White minority; 223 Mixed; 91 Black-Caribbean; 87 Black African; 197 Indian; 229 Pakistani; 79 Bangladeshi; [16 Chinese, 131 Other]	Reported by child's parent	Common mental health problems: clinical rating.	White minority —**; Mixed race —**; Black Caribbean —**, Black African ↓**; Indian ↓**; Pakistani —**; Bangladeshi —**	F
Berry 2006 [392]	2	Birmingham and Leicester†, ??	Cross-sectional: 240 children aged 13-18: 120 White British; 120 Indian	Reported by child	Behavioural problems: self-report	Indian ↓*	A, C, ?H,

Study ref.	Pub. type	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Fuller 2006 [393]	2	England, 2004	Cross-sectional: 1120 children aged 4-15: 212 Irish; 168 Black Caribbean; 148 Black African; 194 Indian; 185 Pakistani; 116 Bangladeshi; 97 Chinese. Compared with a nationally representative general population of 5882 children from 2002	Reported by child's parent	Common mental health problems: parent	White Irish —; Black Caribbean —; Black African ↓/—; Indian —; Pakistani —; Bangladeshi —; Chinese —/↓.	A, F, J (age), K
Scott 2006 [394]	2	London, 2001-04	Cross-sectional: 174 children aged 4-7: 42 White British; 32 Black Caribbean; 79 Black African; [21 Other]	Reported by child's parent	Common mental health problems (various measures): parent, teacher, directly observed behaviour	Black Caribbean —/↓**; Black African —/↓**	I, K
Wardle 2006 [628]	1	London, 1999-2004	Prospective cohort study N=5704 children aged 11/12 at baseline and followed up to age 14/15: Over the full study period: 3324 White; 1482 Black/mixed black (1289 Black, 193 mixed black); 627 South Asian/mixed Asian (557 Asian, 70 mixed Asian); [271 Other]	Reported by child	Common mental health problems: self-report	Black/mixed Black ↓***; South Asian/Mixed Asian ↓***	A, F, L
Laurens 2007 [422]	3	London, ??	Cross-sectional: 595 children aged 9-12: 200 White British; 45 White Minority; 102 Black Caribbean; 155 Black African; 45 South Asian/Chinese; [43 Other]	Reported by child's parent	Psychotic-like experiences: self-report	White minority —; Black Caribbean ↑; Black African —; South Asian/Chinese ↓	A, ?C, F, H, K
Maynard 2009 [413]; Maynard 2007 [414]	3; 1	London, 2003†	Cross-sectional: 4635 children aged 11-13: 1224 White British; 297 Mixed White British/Black Caribbean; 926 Black Caribbean; 609 Nigerian or Ghanaian; 464 Other African; 494 Indian; 621 Pakistani/Bangladeshi	Reported by child	Common mental health problems: self-report	Mixed Black Caribbean/White —; Black Caribbean —; Nigerian/Ghanaian ↓; Other African ↓; Indian ↓	A
See notes to **Table 14.5**							

Table 14.5: Description of clinic-based studies of child mental health

Study ref.	Publication	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Graham 1967 [395]	1	London, 1963-65	Retrospective case notes analysis: 110 children: 55 White British; 55 Black Caribbean	Not specified	Proportional morbidity for diagnosis	Black Caribbean PM: behavioural —/↑; emotional ↓ mixed emotional/behavioural —	?E, G, L
Taylor 1984 [396]	1	London, ??	Case-control study 100 children aged 8-17: 14 Black Caribbean; 86 non-Black Caribbean	Country of origin	Proportional morbidity for Deliberate self harm	Black Caribbean PM: DSH —	F, G, K
Glover 1989 [397]	1	London, 1980-84	Retrospective case notes analysis: 156 girls aged 15-19: 25 South Asian; 131 non-South Asian	Not specified	Deliberate self harm	South Asian ↑/—	D, F, G, ?H, K
Stern 1990 [398]	1	London, 1987	Retrospective case notes analysis.: 189 children (142 from Tower Hamlets): Tower Hamlets referrals: 18 Bangladeshi; 124 non-Bangladeshi. All referrals: 27 South Asian; 162 non-South Asian	Not specified	1) Referral to CAMHS 2) Proportional morbidity for referral reason	Tower Hamlets Bangladeshi ↓*. All Bangladeshi PM; Behavioural —; emotional —; DSH —; psychosis —	?E, F, G, J (age, sex), K
Jawed 1991 [399]	1	Bolton, 1981-85	Retrospective case notes analysis: 978 children: Referral data for 45 South Asian, 933 non-South Asian (although it is unclear if details were simply unrecorded in some cases). Diagnosis information on 72 White British; 45 South Asian	Not specified	1) Referral to CAMHS 2) Proportional morbidity for diagnosis	South Asian ↓*. South Asian PM: behavioural ↓/—; emotional —; hyperactivity —; somatoform ↑/—; psychosis —	?E, F, G, ?H, J (age), K
McGibben 1992 [400]	1	Coventry, 1982-90	Retrospective case notes analysis: 340 children aged 12-15: 295 White; 45 South Asian	Not specified	Deliberate self harm	South Asian —*	?E, F, G, ?H, K
Goodman 1995 [401]	1	London, 1973-89	Retrospective case notes analysis: 1603 children aged under 18: 292 second generation Black-Caribbean; 1311 comparison group with both parents born in Britain.	Country of origin	Proportional morbidity for diagnosis	Black Caribbean PM: behavioural ↑; emotional ↓; mixed emotional/behavioural —; hyperactivity —; autism ↑; psychosis ↑	F, G

Study ref.	Publication	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Roberts 1995 [402]	1	Bradford, 1987-91	Retrospective case notes analysis.: 2 462 children aged 2-17: Referrals: 184 Punjabi Pakistani Muslims; 2 278 non-Punjabi Pakistani. PM for diagnoses: 184 White British; 184 Punjabi Pakistani	Not specified	1) Referral to CAMHS 2) Proportional morbidity for diagnosis	Pakistani ↓. Pakistani PM: behavioural ↓; anxiety —; adjustment disorders ↑	?E, G, ?H
Goddard 1996 [403]	1	London, 1990-92	Retrospective case notes analysis: 100 children aged 10-17 : 64 White; 28 Black (20 Black Caribbean, 2 Black African, 6 Other Black); [8 Other].	Not specified	Deliberate self harm	Black —	?E, F, G, ?H, K
Kramer 2000 [404]	1	London, 1991-92	Retrospective case notes analysis: 183 children: 102 White (74 British White; 28 White minority); 18 Black; 14 South Asian; [49 Mixed/Other]	Not specified	Referral to CAMHS	Black ↓*; South Asian —*	?E, F, G, K
Jayarajan 2001 [405]	3	Birmingham, 1998	Retrospective case notes analysis: 1068 children: 821 White; 116 Black (60 Black Caribbean, 4 Black African, 52 Black Other); 17 Indian; 71 Pakistani; 17 Bangladeshi; [44 Other]	Other method	Referral to CAMHS	Black ↑*; Indian ↓*; Pakistani ↓*; Bangladeshi ↓*	E, F, G, H, K
Lamb 2002 [406]	1	London, 1997	Retrospective case notes analysis.: 444 children (380 from Tower Hamlets): 218 Bangladeshi (216 from Tower Hamlets); 316 non-Bangladeshi (254 from Tower Hamlets)	Not specified	1) Referral to CAMHS 2) Proportional morbidity for referral reason	Tower Hamlets Bangladeshi ↓*; All Bangladeshi PM: behavioural —; emotional —; hyperactivity —; DSH —; psychosis —	?E, G, K
Willis 2002 [407]	3	Manchester, 1999-2000	Retrospective case notes analysis: 192 children aged 2-17: 141 White; 22 Black; 15 South Asian; [14 Mixed race/Other]	Not specified	Referral to CAMHS	Black —*; South Asian —*	?E, F, G, ?H, K
Bhugra, Thompson 2003 [408]	1	London, 1994-95	Retrospective case notes analysis: 61 children age 10-?19: 46 White British/Irish; 15 South Asian	Reported by child	Deliberate self harm	South Asian —*	F, G, ?H, K
Messent 2003 [409]	1	London, 1997	Retrospective case notes analysis: 627 children aged 0-18: 352 White British; 36 Black Caribbean/Black British; 33 Black African; 206 Bangladeshi	Not specified	Referral to CAMHS	Black Caribbean —*;Black African —*; Bangladeshi ↓*	?E, G, K

Study ref.	Public ation	Setting, date	Study design: study population	Ethnicity assignment	Mental health outcome: informant(s)	Results	Limitations
Minnis 2003 [410]	1	Glasgow, 2000-01†	Retrospective case notes analysis: 219 children: 17 South Asian; 212 non-South Asian	Not specified	Attendance at CAMHS	South Asian ↓**	?E, F, G, ?H, K
Tolmac 2004 [411]	1	London, 2001	Cross-sectional: 110 children aged 13-17: 66 White; 21 Black; 11 Asian; [12 Other]	Not specified	CAMHS in-patient with 1) non-psychotic or 2) psychotic disorder	Non-psychotic disorders: Black ↓*; South Asian ↓*. Psychotic disorders: Black ↑*; South Asian —*	?E, F, G, K (non-psychotic disorders)/L(psychotic disorders)
Hackett 2004 [412]	3	Manchester, 2002-03	Retrospective case notes analysis: 320 children aged 2-17: 205 White; 35 Black; 58 South Asian; [22 Mixed race/Other]	Not specified	Referral to CAMHS	Black ↑*; South Asian ↑*	?E, F, G, ?H, K

Notes to **Table 14.4** and **Table 14.5**: CAMHS=Child and Adolescent Mental Health Services. †=information provided through personal communication with authors. Publication type: (1)=published in peer-reviewed journal, (2)=published outside a peer-review journal (e.g. book), (3)=unpublished report/under submission. Ethnic groups listed in square brackets are not presented in results because of insufficient numbers or because they were of 'Other' ethnicity and therefore ineligible. Results: ↑=evidence at the 5% significance level of more mental health problems relative to White/White British/general population, ↑/—=mixture of evidence of more mental health problems and no difference; —=no evidence of difference; ↓/—=mixture of evidence of fewer mental health problems and no difference; ↓=evidence of fewer mental health problems. '?' indicates the apparent trend in cases where there was evidence of overall ethnic differences, but post hoc tests of specific contrasts were not provided. *=statistical calculations performed by AG using information in the text **=statistical calculations performed by AG using information supplied separately by the authors. ***=unpublished results obtained by AG from authors. Methodological limitations (full details in Box 14.2): A=Questionnaire measure of mental health only. B=Reliance on a single, inappropriate informant C=Non-validated modification of a validated mental health measure. D=Method of assigning ethnicity not described. E=Ethnicity determined by an inferior means. F Ethnicity analysed using meta-level descriptions (e.g. 'South Asian') G Potential for selection bias – clinic-based sampling. H Potential for selection bias – response rates/data completeness <60%. I Potential for information bias – investigator-based ratings made by the study authors without being blinded to ethnicity.. J Differences could be due to confounding by age and sex. K No data presented on socio-economic position. L No adjustment made for reported differences in socio-economic position. '?' indicates insufficient information to know whether a limitation applied.

Table 14.6: Description of population-based studies of common mental disorders (CMD) in adults

Study name	Pub. type	Setting, date	Study design: study population	Ethnicity assignment	Mental health assessment	Results (compared to White/White British)	Limitations
Fourth National Survey of Minority Ethnic Groups [243, 307]	2; 2	England and Wales, 1994	Cross-sectional: 8063 adults aged 16-64: 2654 White British; 213White Minority; 614 Black Caribbean; 988 Indian/African Asian; 584 Pakistani; 289 Bangladeshi; 104Chinese	Self-report	Adult CMD (depression and anxiety): self-report	For anxiety: ↑White minority; ↓Black Caribbean; ↓Indian/African Asian; ↓Pakistani; ↓Bangladeshi; ↓Chinese For depression: —* White minority; ↑*Black Caribbean; — * Indian/African Asian; —*Pakistani; ↓*Bangladesh; ↓* Chinese	H, L
Ethnic Minority Psychiatric Illness Rates in the Community (EMPIRIC) [244]	1	England, 2000	Cross-sectional: 4281 adults aged 16-74: 837 White; 733 Irish; 694 Black Caribbean; 650 Indian; 643 Pakistani; 724 Bangladeshi.	Self-report	Adult CMD (depression and anxiety): self-report	Among men: Irish ↑; Black Caribbean — ; Indian — ; Pakistani — ; Bangladeshi — ; Among women: Irish —; Black Caribbean — ; Indian — ; Pakistani ↑ ; Bangladeshi ↓	
National Survey of Psychiatric Morbidity [424]	1	Great Britain, 2000	Cross-sectional: 10 108 adults aged 16–64: 9272 White; 168 Black Caribbean/Black African; 233 South Asian; and 81 Other.	Self-report	Adult CMD (depression and anxiety): self-report	Black Caribbean/Black African — ;South Asian —; and Other —	F
See notes to **Table 14.5**							

14.2 Appendix to Chapter 5: Supplementary information on mental health measures used in B-CAMHS

14.2.1 SDQ items and factor structure

Table 14.7: Summary of which items are proposed to load onto alternative factors in SDQ

	Total difficulty score	**Proposed five-factor structure**					**Alternative three-factor structure**		
		Emotional	Peer problems	Behavioural	Hyper-activity	Pro-social	Internalising	Externalising	Pro-social
Somatic	✓	✓					✓		
Worries	✓	✓					✓		
Unhappy	✓	✓					✓		
Clingy	✓	✓					✓		
Fears	✓	✓					✓		
Solitary	✓		✓				✓		
Good friend*	✓		✓				✓		
Popular*	✓		✓				✓		
Bullied	✓		✓				✓		
Best with adults†	✓		✓				✓		
Tempers	✓			✓				✓	
Obedient*	✓			✓				✓	
Fights	✓			✓				✓	
Lies	✓			✓				✓	
Steals	✓			✓				✓	
Restless	✓				✓			✓	
Fidgety	✓				✓			✓	
Distractible	✓				✓			✓	
Reflective*	✓				✓			✓	
Persistent*	✓				✓			✓	
Considerate*						✓			✓
Shares*						✓			✓
Caring*						✓			✓
Kind to kids*						✓			✓
Helps out*						✓			✓

* Indicates positively worded strengths which are reversed scored. † This item is technically neutral but is scored as a difficulty because it is indicative of peer problems. For full questions for each item, see annotated copy of the parent SDQ, p.418

14.2.2 Psychometric properties of the SDQ and DAWBA

For definitions of reliability and validity concepts and measures, see Appendix 1, Table 13.1 and Table 13.2.

SDQ

Table 14.8: Summary of reliability and validity for the SDQ total difficulty score in Britain

RELIABILITY	
Test-retest reliability	Two very small samples (N<40) gave point estimates of test-retest reliability of an intraclass correlation coefficient of 0.85 for parents [452] and Pearson's correlation coefficient 0.65/0.83 (for paper administration/computer administration respectively) for children [453]. [In B-CAMHS99, Pearson correlation coefficients over six months were 0.72 for parent TDS (N=2091), 0.80 for teacher TDS (N=796) and 0.62 for child TDS (N=781). This is really a test for medium/long-term stability, but Goodman *et al.* [448] argue that this provides a lower bound on test-retest reliability.]
Inter-rater reliability	Cross-informant Pearson correlations for B-CAMHS99 were 0.46 for parent-teacher correlation, 0.48 for parent-child correlation and 0.33 for teacher-child correlation [448]. For comparison, a meta-analysis of cross-informant correlations for a range of mental health measures in children aged 1-18 years calculated mean inter-rater Pearson correlations of 0.27 between parents and teachers (41 samples), 0.25 between parents and children (14 samples) and 0.20 between teachers and children (21 samples) [226]. Previous comparisons had also indicated that parent-teacher correlation was at least as high for the SDQ as the Rutter [194].
Internal consistency	In B-CAMHS99, Cronbach α of 0.82 for parent TDS, 0.87 for teacher TDS and 0.80 for child TDS [448].
VALIDITY	
Content validity	Subscales were chosen based on the main nosological categories (emotional, behavioural and hyperactivity), as were individual items within this (e.g. two inattention, two hyperactivity and one impulsiveness in hyperactivity subscale) [194].
Criterion validity	Not strictly applicable in child mental health, due to the lack of an absolute gold standard (see Section 2.4, Chapter 2)
Construct validity	
a) Convergent and discriminant validity	**Correlation of SDQ TDS with established dimensional measures**. High correlation (Pearson coefficients 0.78-0.92) of SDQ TDS with the parent and teacher Rutter and parent ASEBA [194, 451]. SDQ functions at least as well as these in identifying high-risk groups. **Sensitivity and specificity of SDQ caseness with face-to-face clinical diagnosis.** 101 consecutive new referrals aged 4-16 to a London clinic, divided using multi-informant SDQ algorithm into 'unlikely/possible' vs. 'probable' cases. Sensitivity was 90% for conduct disorders, 81% for emotional disorders and 89% for hyperactivity disorders. Note also that all but one of the 'false negatives' are partial, corresponding to instances in which the child was rated as having a 'possible' disorder. Specificity was 47% for conduct disorders, 81% for emotional disorders and 78% for hyperactivity disorders. The SDQ algorithm therefore appears to be a sensitive but over-inclusive in detecting disorders [454].

	Sensitivity and specificity of high SDQ scores with DAWBA diagnosis. In B-CAMHS99, children with scores above the 90th percentile had an odds ratio of receiving a DAWBA diagnosis of over 15 for parents and teachers, and an odds ratio of over six for children [448]. Compared to any DAWBA diagnosis, the multi-informant SDQ had sensitivity of 63% and specificity of 95% (with the sensitivity for particular diagnoses being 76% for conduct disorder; 75% for DSM-IV ADHD; 75% for any depressive disorder; and 51% for any anxiety disorder). In an equivalent analyses performed in separate sample of looked-after children, the sensitivity was 85% and the specificity 80% [455]. **Relationship between baseline SDQ TDS and DAWBA diagnosis at baseline and follow-up.** In B-CAMHS99 and B-CAMHS04, each one-point increase in the TDS was associated with an increased prevalence of DAWBA diagnosis. This was true across the full range of the TDS in all three informants. The same relationship was observed in all informants when using DAWBA diagnosis at three-year follow-up [456].
b) Group differentiation, hypothesis testing and prognosis	**Ability of the SDQ TDS to distinguish between psychiatric clinic samples and non-clinic samples.** In Receiver Operating Characteristics (ROC) analyses, the TDS yielded and area under the curve of 0.87 for parent total SDQ and 0.85 for teacher SDQ in one comparison [194], and 0.93 for parent total SDQ in a second comparison [451]. In both comparisons, the SDQ seemed to function at least as well as the Rutter and the parent ASEBA. **Ability of parent, teacher and child SDQ 'caseness' to distinguish between psychiatric clinic samples and non-clinic samples.** Chance-corrected kappas were 0.68 for the parent SDQ, 0.56 for teacher SDQ and 0.40 for the child SDQ [452]. **Risk factor profiles for the SDQ compared to the DAWBA.** Unpublished analyses of the B-CAMHS99 baseline survey indicate that similar or identical risk factors predict mean SDQ score as predict DAWBA diagnoses [458]. **Risk factor profiles for high, low and mean SDQ scores.** In B-CAMHS99, the same risk factors that predict change in SDQ score across the entire range, also predict it in children one standard deviation above and one standard deviation below the mean [457]. This provides some indirect evidence as to the validity of the SDQ as a measure of child mental health across its full range.

Table 14.9: Summary of reliability and validity for the SDQ subscales in Britain

RELIABILITY	
Test-retest reliability	Test-retest reliability has not been assessed for the individual subscales although, as for the TDS, there is data from B-CAMHS99 about stability over six months [448]. Again, stability was highest for teachers (range 0.65-0.82), followed by parents (range 0.57-0.72) and lowest for child-report (range 0.51-0.60). Stability was highest for the hyperactivity subscale.
Inter-rater reliability	In B-CAMHS99, cross-informant subscale correlations gave Pearson's coefficients of between 0.25-0.48 for parent-teacher agreement, 0.30-0.44 for parent-child and 0.21-0.32 for teacher-child agreement [448]. These are in most cases higher than the meta-analytic mean correlations of 0.27 between parents and teachers, 0.25 between parents and children and 0.20 between teachers and children [226].
Internal consistency	Excluding the peer problems subscales, the Cronbach α in B-CAMHS99 were 0.63-0.77 for the parent subscales, 0.74-0.88 for the teacher subscales and 0.60-0.67 for the child subscales. For the peer problems subscales the Cronbach α were lower, being 0.57 for parents, 0.70 for teachers and 0.41 for children [448].
VALIDITY	
Content validity	Subscales were chosen based on main nosological categories (emotional, behavioural and hyperactivity), as were individual items within this (e.g. two inattention, two hyperactivity and one impulsiveness in hyperactivity subscale) [194].
Criterion validity	Not strictly applicable in child mental health, due to the lack of an absolute gold standard (see Section 2.4, Chapter 2)
Construct validity	
a) Convergent and discriminant validity	**SDQ subscales compared to DAWBA diagnosis.** In B-CAMHS99, children with a DAWBA diagnoses for emotional, behavioural or hyperactivity disorders scored at least one standard deviation higher than the mean on the corresponding SDQ subscale in all three informants [459]. They also scored less highly (often much less highly) on non-corresponding SDQ subscales than on the expected subscale. The only exception of teacher-reported emotional problems, which performed little better than teacher-reported behavioural problems in predicting emotional disorder. **SDQ subscales compared to PACS diagnosis.** The Parental Account of Child Symptoms [PACS: 460] is an investigator-based, semi-structured interview which can be used to generate scores for emotional, behavioural and hyperactivity problems. This was administered to children in psychiatric clinic and non-clinic populations, and showed correlations of 0.64 with the SDQ behavioural scale, 0.53 with the emotional scale and 0.43 with the hyperactivity scale [451]. All of these were at least as good as the parent ASEBA, and in the case of the hyperactivity subscale were substantially better,.
b) Group differentiation, hypothesis testing and prognosis	**Pattern of cross-scale correlations**. In B-CAMHS99, averaging across parents, teachers and children produced cross-scale Pearson correlations of 0.28 for emotional-behavioural, 0.27 for emotional-hyperactivity and 0.55 for behavioural-hyperactivity. This confirms the expectation of substantially higher correlation between the two subscales of externalising symptoms than of either externalising subscale with emotional (internalising) symptoms. The correlation for the SDQ's internal and externalising subscales is lower than that seen for other measures such as the ASEBA, which may suggest that the scales in the SDQ are 'purer' [448, 451].
c) Factor analysis	**Principal component analysis:** In five-factor principal component analyses for B-CAMHS99, all 25 items loaded onto their expected scale in all three informants, with loadings of >0.4 for 72/75 cases [448]. 18/75 items loaded onto other factors, in 22 loadings, most of these being in the teacher (7 items) and child (8 items) SDQs. In most cases, these cross-loadings were small (<0.4 for 18/22 loadings) and were smaller than the loading onto the expected factor (16/18 items). 9/22 of the unexpected loadings (5/6 of the loadings of >0.4) were strengths which loaded onto the 'prosocial' factor.

For a definition of reliability and validity concepts and measures, see Appendix 1, Table 13.1 and Table 13.2.

Table 14.10: Summary of reliability and validity for DAWBA diagnoses in Britain

RELIABILITY	
Test-retest reliability	Not assessed
Inter-rater reliability	
a) Agreement between informants (parents, teachers, children)	Not assessed; instead information from all informants is synthesised into a single diagnosis.
b) Agreement between clinicians	In B-CAMHS99, 500 children were randomly selected and independently rated by different clinicians on the B-CAMHS team. The weighted kappa (κ) statistics for chance-corrected agreement was 0.86 for any disorder, 0.57 for internalizing disorder and 0.98 for externalising disorder [416].
VALIDITY	
Content validity	Based closely on the DSM-IV/ICD-10 criteria
Criterion validity	Not strictly applicable in child mental health, due to the lack of an absolute gold standard (see Section 2.4, Chapter 2)
Construct validity	
a) Convergent and discriminant validity	**Comparison of DAWBA diagnoses with case notes diagnoses**. DAWBA diagnoses were compared with diagnoses from a case notes review for 39 children in a mental health clinic. Review of the clinic notes suggested 30 diagnoses were 'definitely present', and the DAWBA identified the same diagnosis in 28 instances (93%). The DAWBA also identified a further 28 disorders, however, 17 of which had been judged only 'possibly present' and 11 which were not suggested by the case notes review. This resulted in a final Kendall tau correlation of 0.56 [459]. For comparison, a meta-analysis of agreement between clinicians and standardised diagnostic interviews produced a chance-corrected kappa of 0.21 [471]. Moreover, it should be noted that 19/28 false positives were of comorbid diagnoses, where the DAWBA agreed on the principal diagnosis in the case notes but also diagnosed additional disorders. Finally, interpretation is complicated by inadequacies in the case notes, which often lacked detail and were made by a variety of professionals of different levels of seniority. Interpretation is also complicated by the small sample size.
b) Group differentiation, hypothesis testing and prognosis	**Differentiation of clinic and community samples**. 36/39 children in a clinic (92%) were given a diagnosis, compared to 52/491 (11%) in the community [459]. [NB this study also demonstrated that the skip rules were, as intended, functioning to reduce interview length substantially without seriously compromising interview quality.] **Correlates and risk factor profiles of general population children with a diagnosis**. 52 children in a community sample who received DAWBA diagnoses were compared with 439 children who did not. Children with a diagnosis had odds ratios of between 8 and 27 for five selected indicators of mental health problems (parents say there is a problem; teachers say there is a problem; child says there is a problem; the child receives mental health care; the child receives extra help in schools). Children with a DAWBA diagnosis also showed poorer prognosis over 4-6 months (as judged using the SDQ) than children with high scores at baseline (top 20%) but not receiving a diagnosis [459]. **Reasonable prevalence estimates in B-CAMHS**. At 10%, the prevalence estimates for mental disorders generated in the B-CAMHS surveys are in line with those produced by other methods [4]. The same is true of the prevalences of different types of disorder [2-3].

	Risk factor profiles in B-CAMHS. Risk factor profiles of children with emotional, behavioural and hyperactivity disorders differ markedly from the general population. They also differ from each other in ways in line with the literature [3, 417]; see also Section 5.4.3, Chapter 5. **Prognosis to three-year follow-up in B-CAMHS**. DAWBA diagnosis in B-CAMHS strong predictive for DAWBA re-diagnosis at follow-up. It is also strongly predictive of mental health service use and of adverse non-mental health outcomes such as exclusion [48, 445]

For a definition of reliability and validity concepts and measures, see Appendix 1, Table 13.1 and Table 13.2.

14.2.3 Supplementary information on creating and evaluating the DAWBA bands

Creating the DAWBA bands

Table 14.11 provides details of the predicted probability categories which the original B-CAMHS team created for each individual disorder. It further describes how I then used these to create the emotional, behavioural, hyperactivity and 'any common disorder' DAWBA bands. It also describes the underlying predicted probability categories for each individual disorder.

Table 14.11: Constituent diagnoses for each of the DAWBA bands, and underlying predicted probability categories.

	DAWBA band (highest of any constituted individual disorder)		Individual disorders for which probability categories calculated	Probability categories for individual disorders†					
Parent	Any common disorder	Emotional disorders	Separation anxiety		0.5%	3%	15%		>70%
			Specific phobia	<0.1%	0.5%		15%	50%	
			Social phobia	<0.1%	0.5%	3%	15%	50%	
			PTSD [2004 only]	<0.1%	0.5%	3%	15%	50%	
			Panic	<0.1%		3%	15%		
			Agoraphobia	<0.1%		3%	15%		
			OCD	<0.1%	0.5%	3%	15%	50%	
			GAD		0.5%	3%	15%	50%	
			Depression	<0.1%	0.5%	3%	15%	50%	>70%
		Behavioural disorders	ODD		0.5%	3%	15%	50%	>70%
			Conduct disorder		0.5%	3%	15%	50%	>70%
		Hyperactivity	ADHD	<0.1%	0.5%	3%	15%	50%	>70%
Teacher	Any common disorder	Emotional disorder	'Any emotional disorder'	<0.1%	0.5%	3%	15%		
		Behavioural disorders	ODD		0.5%	3%	15%	50%	>70%
			Conduct disorder		0.5%	3%	15%	50%	>70%
		Hyperactivity	ADHD	<0.1%	0.5%	3%	15%	50%	
Child	Any common disorder	Emotional disorder	Separation anxiety		0.5%	3%	15%		>70%
			Specific phobia	<0.1%	0.5%		15%	50%	
			Social phobia	<0.1%	0.5%	3%	15%	50%	
			PTSD [2004 only]	<0.1%	0.5%	3%	15%	50%	
			Panic	<0.1%		3%	15%	50%	
			Agoraphobia	<0.1%		3%	15%	50%	
			OCD	<0.1%	0.5%	3%	15%	50%	
			GAD		0.5%	3%	15%	50%	
			Depression	<0.1%	0.5%	3%	15%	50%	>70%
		Behavioural disorders	Conduct disorder		0.5%	3%	15%	50%	

PTSD=Post-traumatic stress disorder; OCD=Obsessive compulsive disorder; GAD=generalised anxiety disorder; ODD=oppositional defiant disorder; ADHD=Attention Deficit Hyperactivity Disorder.

† Some cells are empty because no set of responses corresponded empirically to that nominal probability of receiving a diagnosis. For example a '<0.1% probability' category could not be created for separation anxiety by parent report because even when the parent DAWBA provided no indication of a separation anxiety disorder, nonetheless the observed rate of diagnosis was closer to 0.5% than 0.1% (in such cases, diagnosis would be based upon a convincing account of separation anxiety in the parent open-ended transcript or in the child DAWBA).

<u>Evaluating the DAWBA bands</u>

The DAWBA bands as ordered categorical variables

My intention in creating the DAWBA bands was to generate informant-specific, ordered-categorical variables of emotional, behavioural and hyperactivity disorders. To assess whether the DAWBA bands were indeed ordered categorical, I examined whether children with higher DAWBA bands had higher SDQ scores and a higher prevalence of DAWBA diagnosis. Using the baseline SDQ scores and DAWBA diagnoses would be circular because the SDQ scores form part of the skip rule which determines which sections of the

DAWBA are administered, and because the structured sections of the DAWBA form part of the information used to assign diagnoses. I therefore compared the DAWBA bands at baseline with the SDQ scores and DAWBA diagnoses at three-year follow-up.

Table 14.12 presents the results of these analyses for the ‘any common disorder’ DAWBA bands; the results were very similar for the separate emotional, behavioural and hyperactive DAWBA bands. By all measures there is strong evidence ($p<0.001$) of poorer mental health for higher DAWBA bands. Differences between categories are large at every level, almost always with at least a 2-3 fold increase in prevalence or a mean difference of 2-3 SDQ points. This includes mental health outcomes reported by independent informants (e.g. mean teacher SDQ compared to parent DAWBA band). This therefore provides good evidence that the DAWBA bands are ordered categorical measures of mental health problems.

Table 14.12: DAWBA diagnosis and SDQ at follow-up by level at baseline on the DAWBA band for any common mental disorder

	DAWBA band	Prevalence disorder, all children	Prevalence disorder, no disorder at baseline	Mean parent TDS	Mean teacher TDS	Mean child TDS
Parent	**1 (lowest risk)**	3.3%	3.2%	4.4	4.2	8.1
	2	6.1%	5.4%	7.7	6.0	10.1
	3	20.5%	15.7%	12.6	9.2	12.5
	4 (highest risk)	53.8%	30.0%	19.1	13.4	15.4
Teacher	**1 (lowest risk)**	4.7%	3.8%	6.2	3.2	8.9
	2	10.9%	7.6%	8.6	8.1	10.7
	3	24.2%	13.8%	12.7	15.2	14.0
	4 (highest risk)	41.1%	30.5%	14.7	19.1	13.6
Child	**1 (lowest risk)**	3.8%	3.2%	5.6	4.2	7.2
	2	9.6%	6.8%	7.5	5.7	10.7
	3	20.3%	14.9%	10.4	8.3	13.7
	4 (highest risk)	32.8%	17.0%	13.0	10.9	16.1

All differences/gradients in mental health problems significant at $p<0.001$.

Construct validity of the DAWBA bands

I then assessed the convergent and discriminant validity of the parent, teacher and child DAWBA bands for emotional, behavioural and hyperactivity disorders. I did so by performing a series of logistic regression analyses in which the outcomes were follow-up diagnoses of emotional, behavioural and hyperactivity disorders. The explanatory variables were the baseline emotional, behavioural and hyperactivity DAWBA bands from a particular informant. *A priori*, I expected a diagnosis of emotional disorder to be most strongly predicted by the emotional DAWBA band, and so on.

As Table 14.13 shows, this was indeed the case: the expected DAWBA band always showed a strong association with the corresponding DAWBA diagnosis at follow-up (OR 2.09-3.61), while other subscales were either not independently predictive or else showed substantially smaller OR. Sensitivity analyses which re-ran the regression models adjusting for comorbidity (i.e. adjusting for DAWBA diagnosis for behavioural and hyperactivity disorder when predicting to emotional disorders etc.) produced an even cleaner picture of convergent and discriminant validity. This therefore provides evidence of the construct validity of the emotional, behavioural and hyperactivity DAWBA bands.

Table 14.13: Logistic regression analyses predicting DAWBA diagnosis at follow-up from DAWBA bands at baseline

		Emotional disorder at follow-up (OR and 95% CI)	**Behavioural disorder at follow-up (OR and 95% CI)**	**Hyperactivity disorder at follow-up (OR and 95% CI)**
Parent (N=7825)	**Emotional DAWBA band**	2.19 (1.96, 2.45)***	1.09 (0.95, 1.24)	1.00 (0.82, 1.22)
	Behavioural DAWBA band	1.30 (1.10, 1.53)**	3.60 (3.03, 4.28)***	2.27 (1.71, 3.01)***
	Hyperactivity DAWBA band	1.14 (0.96, 1.35)	1.46 (1.26, 1.70)***	3.21 (2.48, 4.15)***
Teacher (N=5972)	**Emotional DAWBA band**	2.09 (1.58, 2.76)***	1.37 (1.06, 1.78)*	0.99 (0.66, 1.49)
	Behavioural DAWBA band	1.57 (1.30, 1.90)***	2.31 (1.95, 2.73)***	1.57 (1.21, 2.04)**
	Hyperactivity DAWBA band	0.84 (0.68, 1.04)	1.36 (1.14, 1.63)**	2.87 (2.15, 3.83)***
Child (N= 3242)	**Emotional DAWBA band**	2.07 (1.77, 2.42)***	1.28 (1.06, 1.56)*	–
	Behavioural DAWBA band	1.13 (0.90, 1.41)	2.76 (2.26, 3.36)***	–

*p<0.05, **p<0.01, ***p<0.001. Cells shaded grey indicate the subscales *a priori* expected to be the strongest predictor. Note that no DAWBA band exists for child hyperactivity.

14.3 Appendix to Chapter 6: Imputing ethnicity using the name-matching software Onomap

The Child Benefit Centre provided ONS with the child's first name and surname for all children who participated in B-CAMHS and three-quarters of those who did not. This allowed me to use name-matching software to learn something about the likely ethnicity of non-participants.

I used the name-matching software Onomap (www.onomap.org). This was developed for use in the UK, and uses forenames and surnames to group together names with a common cultural, ethnic and/or linguistic origin [481]. These can then be converted into 2001 UK census categories.

Application of Onomap to the B-CAMHS dataset

Onomap was validated using data from British adults [481], and no validation has been conducted using children. I therefore evaluated the performance of Onomap in B-CAMHS by comparing it to parent-reported ethnicity for those children who did participate. I did so through cross-tabulations (Table 14.14) and by calculating the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of each Onomap category relative to parent-reported ethnicity (Table 14.15).

The results in Table 14.15 indicate that Onomap functioned fairly well at identifying Indians, with reasonably high sensitivity (71%) and PPV (82.8%) although underestimating somewhat the total number (360 by Onomap vs. 419 by parent-report). The most common misclassification was to class Indians as Pakistani (9.1% of Indians) or Other (13.1%). Specificity and NPV are high for Indians, as for all non-White ethnic groups. Onomap also performed quite well in identifying Pakistani and Chinese children, but sensitivity is poor for Bangladeshi (48%) and Black African (37%) names, and the number of children in each category is substantially underestimated. For 'Other' and Black Caribbean names, Onomap performed very poorly indeed, identifying only 3/196 Black Caribbean children correctly. This reflects the close similarity between the names of Black Caribbean and White children.

This preliminary evaluation therefore suggests that Onomap may be useful in identifying Whites and Indians but that it has substantially poorer performance for some other groups. I therefore proceed cautiously in Section 6.3, Chapter 6 when using Onomap to investigate participation rates for different ethnic groups, and provide further evaluation of its adequacy for this purpose.

Table 14.14: Comparison of Onomap and parent-reported ethnicity in B-CAMHS

		Child's ethnicity by parent self-report								
		White	**Black-Caribbean**	**Black-African**	**Indian**	**Pakistani**	**Bangladeshi**	**Chinese**	**Other**	**ALL**
Onomap ethnicity	**White**	**16295**	169	55	18	3	4	14	393	16951
	Black-Caribbean	3	**3**	0	0	0	0	0	5	11
	Black African	19	3	**60**	4	3	0	0	25	114
	Indian	9	1	2	**298**	13	1	0	36	360
	Pakistani	15	6	9	38	**244**	25	0	45	382
	Bangladeshi	3	1	1	3	11	**59**	0	7	85
	Chinese	5	1	1	0	0	0	**35**	13	55
	Other	99	12	33	58	102	33	1	**114**	452
	ALL	16448	196	161	419	376	122	50	638	18410

1 B-CAMHS participant was missing data on Onomap ethnicity and 4 on parent-reported ethnicity

Table 14.15: Predictive performance of Onomap categories against parent-reported ethnicity

Ethnic group	**Number by parent-reported ethnicity**	**Number by Onomap imputed ethnicity**	**Onomap sensitivity**	**Onomap specificity**	**Onomap positive predictive value (PPV)**	**Onomap negative predictive value (NPV)**
White	16448	16951	99.1%	66.6%	96.1%	89.5%
Black-Caribbean	196	11	1.5%	99.96%	27.3%	99.0%
Black African	161	114	37.3%	99.7%	52.6%	99.5%
Indian	419	360	71.1%	99.7%	82.8%	99.3%
Pakistani	376	382	64.9%	99.2%	63.9%	99.3%
Bangladeshi	122	85	48.4%	99.9%	69.4%	99.7%
Chinese	50	55	70.0%	99.9%	63.6%	99.9%
Other	638	452	17.9%	98.1%	25.2%	97.1%

14.4 Appendix to Chapter 7: Supplementary data regarding descriptive analyses of mental health outcomes of Indians and Whites

14.4.1 Individual diagnoses in Indians and Whites

Table 14.16: Number of children receiving each individual clinical diagnosis among White (N=16 435) and Indian (N=419) children

	Diagnosis	White	Indian
Anxiety disorders	Any anxiety disorder	589	5
	• Separation anxiety	175	0
	• Specific phobia	152	1
	• Social phobia	54	0
	• Panic disorder	26	1
	• Agoraphobia	17	0
	• Post-traumatic stress disorder	26	0
	• Obsessive compulsive disorder	30	0
	• Generalised anxiety disorder	119	3
	• Other anxiety-NOS	124	1
	Among children with anxiety disorders, no diagnosis showed evidence (p<0.05) of being proportionally over- or under-representation in Indian children		
Depressive disorders	Any depressive disorder	146	6
	• Major depression	110	5
	• Other depression-NOS	36	1
	Among children with depressive disorders, no diagnosis showed evidence (p<0.05) of being proportionally over- or under-representation in Indian children		
Behavioural disorders	Any behavioural disorder	858	5
	• Oppositional defiant disorder	524	3
	• Conduct disorder	303	2
	• Other behavioural disorder-NOS	139	0
	Among children with behavioural disorders, no diagnosis showed evidence (p<0.05) of being proportionally over- or under-representation in Indian children		
Hyperactive disorders	Any hyperactive disorder	386	1
	• ADHD, combined type	254	0
	• ADHD, inattentive type	103	1
	• ADHD, hyperactive-impulsive type	29	0
	Among children with hyperactive disorders, no diagnosis showed evidence (p<0.05) of being proportionally over- or under-representation in Indian children		
Less common disorders	All less common disorders	148	1
	• Pervasive developmental disorder	88	1
	• Tic disorder	31	0
	• Eating disorder	22	0
	• Other disorder-NOS	13	0
	Among children with less common disorders, no diagnosis showed evidence (p<0.05) of being proportionally over- or under-representation in Indian children		

ADHD=attention deficit/hyperactivity disorder, NOS=Not Otherwise Specified. All assessments of proportional over- or under-representation were calculated Fisher's exact chi2 tests. This was done without adjusting for survey design, as exact chi2 tests is not an option supported by the Stata svyset commands.

14.4.2 Interactions between ethnicity and age/gender in predicting SDQ scores

Table 14.17: Comparison of mean parent SDQ scores (N=16 386 for White parents, N= 389 for Indian parents) with additional information on interactions

	Group or subgroup	White mean	Indian mean	Regression coefficient & 95% CI for White (vs. Indian) ethnicity	Interaction with age or gender ($p \leq 0.01$)?
20-ITEM SCALE					
Total difficulty score (neg)	All	8.25	7.58	0.63 (-0.03, 1.29)	*Gender (p=0.001)*
	Boys	*8.93*	*7.48*	*1.43 (0.54, 2.32)***	
	Girls	*7.54*	*7.69*	*-0.21 (-0.96, 0.54)*	
10-ITEM SCALES					
Internalising subscale (neg)	All	3.31	3.63	-0.33 (-0.74, 0.09)	*Gender (p=0.006)*
	Boys	*3.33*	*3.24*	*0.10 (-0.41, 0.61)*	
	Girls	*3.30*	*4.05*	*-0.77 (-1.31, -0.24)***	
Externalising subscale (neg)	All	4.93	3.95	0.96 (0.64, 1.27)***	
5-ITEM SCALES					
Emotional subscale (neg)	All	1.88	1.96	-0.08 (-0.33, 0.18)	
Peer problems subscale (neg)	All	1.43	1.68	-0.25 (-0.45, -0.04)*	*Gender (p=0.001)*
	Boys	*1.53*	*1.51*	*0.02 (-0.24, 0.29)*	
	Girls	*1.33*	*1.85*	*-0.53 (-0.78, -0.28)****	
Behavioural subscale (neg)	All	1.56	1.28	0.27 (0.10, 0.44)**	
Hyperactivity subscale (neg)	All	3.38	2.67	0.69 (0.47, 0.90)***	*Gender (p=0.003)*
	Boys	*3.91*	*2.86*	*1.03 (0.69, 1.37)****	
	Girls	*2.83*	*2.47*	*0.33 (0.03, 0.62)**	
Prosocial subscale (pos)	All	8.69	8.97	-0.27 (-0.44, -0.10)**	

*p<0.05, **p<0.01, ***p<0.001. (pos)=positive scale; a higher score is more favourable; (neg)=negative scale; a higher score is less favourable. All regression coefficients generated through linear regression, and adjust for age, gender and survey year.

Table 14.18: Comparison of mean teacher SDQ scores (N=12 796 for White teachers, N=302 for Indian teachers) with additional information on interactions

	Group or subgroup	White mean	Indian mean	Regression coefficient & 95% CI for White (vs. Indian) ethnicity	Interaction with age or gender ($p\leq0.01$)?
20-ITEM SCALE					
Total difficulty score (neg)	All	6.50	5.16	1.34 (0.64, 2.03)***	
10-ITEM SCALES					
Internalising subscale (neg)	All	2.82	2.52	0.30 (-0.16, 0.76)	
Externalising subscale (neg)	All	3.68	2.64	1.04 (0.69, 1.39)***	
5-ITEM SCALES					
Emotional subscale (neg)	All	1.49	1.14	0.35 (0.09, 0.62)**	
Peer problems subscale (neg)	All	1.34	1.39	-0.06 (-0.30, 0.18)	
Behavioural subscale (neg)	All	0.89	0.49	0.39 (0.26, 0.53)***	Age (p=0.001)
	5-8 yrs	*0.86*	*0.70*	*0.15 (-0.11, 0.41)*	
	9-12 yrs	*0.88*	*0.48*	*0.39 (0.19, 0.59)****	
	13-16 yrs	*0.95*	*0.25*	*0.73 (0.59, 0.87)****	
Hyperactivity subscale (neg)	All	2.79	2.14	0.64 (0.37, 0.92)***	
Prosocial subscale (pos)	All	7.41	7.76	-0.34 (-0.55, -0.13)**	

*p<0.05, **p<0.01, ***p<0.001. (pos)=positive scale; a higher score is more favourable; (neg)=negative scale; a higher score is less favourable. All regression coefficients generated through linear regression, and adjust for age, gender and survey year.

In no instance was there evidence (p<0.01) of an interaction with the child SDQ for any subscale.

14.5 Appendix to Chapter 8: Supplementary information on SDQ psychometric properties

14.5.1 Example of MPlus syntax for confirmatory factor analysis of hypothesised SDQ factor structure

```
TITLE: PARENT CFA of hypothesised factor structure

DATA:
    FILE IS [name_of_input_file] ;

VARIABLE:
    NAMES ARE studyno weight strata cluster
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25;

    USEVARIABLES ARE
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25;

   CATEGORICAL ARE
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25;

    missing are all  (999);
    stratification is strata;
    cluster are cluster;
    weight is weight ;
    idvar is studyno;

ANALYSIS:
    TYPE is complex missing h1;
    ESTIMATOR IS wlsmv;
    ITERATIONS = 100000;
    CONVERGENCE = 0.0005;

MODEL:
    emotional by p1 p2 p3 p4 p5;
    peer by p6 p7 p8 p9 p10;
    behavioual by p11 p12 p13 p14 p15;
    hyperactivity by p16 p17 p18 p19 p20;
    prosocial by  p21 p22 p23 p24 p25;

    tds by emotional peer behavioual hyperactivity;

OUTPUT:
    SAMPSTAT STANDARDIZED;
```

Key to on variable names: studyno=unique individual ID; weight=individual probability weight; strata= strata ID; cluster=cluster ID; p1-p5=five SDQ emotional problems items; p6-p10=five SDQ peer problems items; p11-p15=five SDQ behavioural problems items; p16-p10=five SDQ hyperactivity items; p21-p25=five SDQ prosocial items.

14.5.2 Full details of confirmatory factor analyses in the full B-CAMHS dataset

Table 14.19: Full summary of additional confirmatory factor analyses on 25 items from SDQ using the full B-CAMHS dataset[46]

Informant	Factors	Model structure	CFI	TLI	RMSEA	Acceptable fit by all 3 indices
Parent (N= 18 222)	Int/ext/pro	First order	0.825	0.900	0.071	
	Int/ext/pro	General-specific	0.885	0.934	0.057	
	Int/ext/pro	Second order	0.825	0.900	0.071	
	Emo/peer/behav/hyp/pro	First order	0.879	0.933	0.058	
	Emo/peer/behav/hyp/pro	General-specific	0.861	0.918	0.064	
	Emo/peer/behav/hyp/pro	Second order	0.850	0.915	0.065	
Teacher (N= 14 263)	Int/ext/pro	First order	0.864	0.933	0.114	
	Int/ext/pro	General-specific	0.952	0.978	0.065	√
	Int/ext/pro	Second order [model did not converge]	–	–	–	
	Emo/peer/behav/hyp/pro	First order [model did not converge]	–	–	–	
	Emo/peer/behav/hyp/pro	General-specific	0.920	0.960	0.089	
	Emo/peer/behav/hyp/pro	Second order [model did not converge]	–	–	–	
Child (N=7678)	Int/ext/pro	First order	0.828	0.868	0.066	
	Int/ext/pro	General-specific	0.851	0.883	0.062	
	Int/ext/pro	Second order	0.828	0.868	0.66	
	Emo/peer/behav/hyp/pro	First order	0.850	0.886	0.062	
	Emo/peer/behav/hyp/pro	General-specific	0.822	0.855	0.069	
	Emo/peer/behav/hyp/pro	Second order	0.801	0.846	0.072	

Factor abbreviations: Int=internalising, ext=externalising, pro=prosocial, emo=emotional, peer=peer problems, behav=behavioural, hyp=hyperactivity. Acceptable fit defined as CFI and TLI >0.9, RMSEA<0.08.

[46] Adjusting these models to allow correlations between residual variances resulted in small improvements to the fit of some models, but did not change the overall picture regarding which models provided the best fit. I therefore present results from models without additional adjustments.

Table 14.20: Full summary of additional confirmatory factor analyses on 20 total difficulty items from SDQ using the full B-CAMHS dataset

Informant	No. factors	Model	CFI	TLI	RMSEA	Acceptable fit by all 3 indices
Parent (N= 18 222)	Int/ext	First order	0.857	0.912	0.073	
	Int/ext	General-specific	0.897	0.937	0.062	(√)
	Emo/peer/behav/hyp	First order	0.907	0.945	0.058	√
	Emo/peer/behav/hyp	General-specific	0.882	0.923	0.068	
	Emo/peer/behav/hyp	Second order	0.878	0.926	0.067	
Teacher (N= 14 263)	Int/ext	First order	0.890	0.935	0.113	
	Int/ext	General-specific	0.950	0.974	0.072	√
	Emo/peer/behav/hyp	First order	0.919	0.962	0.086	
	Emo/peer/behav/hyp	General-specific	0.919	0.948	0.102	
	Emo/peer/behav/hyp	Second order	0.908	0.946	0.104	
Child (N=7678)	Int/ext	First order	0.881	0.911	0.059	
	Int/ext	General-specific	0.909	0.928	0.053	√
	Emo/peer/behav/hyp	First order	0.900	0.924	0.054	√
	Emo/peer/behav/hyp	General-specific	0.870	0.896	0.063	
	Emo/peer/behav/hyp	Second order	0.856	0.892	0.065	

Factor abbreviations: Int=internalising, ext=externalising, emo=emotional, peer=peer problems, behav=behavioural, hyp=hyperactivity. Acceptable fit defined as CFI and TLI >0.9, RMSEA<0.08. Second-order two-factor models were not fitted because these do not have enough unknown parameters to be freely estimated.

14.5.3 Full details of exploratory factor analyses in Indians and Whites

Table 14.21: Two-factor EFA for Whites and Indians on the total difficulty items of the parent SDQ

	White (N=16 401)		**Indian (N=389)**	
Item	**"Internalising" factor**	**"Externalising" factor**	**"Internalising" factor**	**"Externalising" factor**
Somatic [i]	**0.44**		**0.48**	
Worries [i]	**0.78**		**0.66**	
Unhappy [i]	**0.75**		**0.66**	
Clingy [i]	**0.57**		**0.63**	
Fears [i]	**0.67**		**0.74**	
Solitary [i]	**0.53**		**0.47**	
Good friend [i] *	**-0.42**		**-0.52**	
Popular [i] *	**-0.46**		**-0.46**	
Bullied [i]	**0.56**		**0.65**	
Best with adults [i]	**0.45**		**0.42**	
Tempers [e]		**0.45**	0.34	0.36
Obedient [e] *		**-0.56**		**-0.41**
Fights [e]		**0.52**		0.39
Lies [e]		**0.56**		**0.43**
Steals [e]		**0.55**		
Restless [e]		**0.75**		**0.73**
Fidgety [e]		**0.75**		**0.64**
Distractible [e]		**0.80**		**0.64**
Reflective [e] *		**-0.68**		**-0.54**
Persistent [e] *		**-0.77**		**-0.57**

[i] indicates item hypothesised to lie on the internalising subscale, [e] on the externalising subscale. Items marked * are positively worded, and therefore expected to load in the reverse direction. Loadings over 0.3 presented, loadings over 0.4 presented in bold.

Table 14.22: Two-factor EFA for Whites and Indians on the teacher SDQ

	White (N=12 865)		**Indian (N=306)**	
Item	**“Internalising” factor**	**“Externalising” factor**	**“Internalising” factor**	**“Externalising” factor**
Somatic [i]	**-0.56**		**-0.52**	
Worries [i]	**-0.83**		**-0.85**	
Unhappy [i]	**-0.75**		**-0.88**	
Clingy [i]	**-0.76**		**-0.65**	
Fears [i]	**-0.90**		**-0.75**	
Solitary [i]	**-0.63**		**-0.79**	
Good friend [i] *	**0.48**	**-0.41**	**0.68**	
Popular [i] *	**0.40**	**-0.55**	**0.52**	-0.37
Bullied [i]	**-0.56**		**-0.61**	
Best with adults [i]	**-0.49**		**-0.58**	
Tempers [e]		**0.68**		**0.49**
Obedient [e] *		**-0.75**		**-0.76**
Fights [e]		**0.79**		**0.81**
Lies [e]		**0.77**		**0.75**
Steals [e]		**0.65**		**0.52**
Restless [e]		**0.96**		**0.88**
Fidgety [e]		**0.95**		**0.86**
Distractible [e]		**0.90**		**0.89**
Reflective [e] *		**-0.82**		**-0.77**
Persistent [e] *		**-0.86**		**-0.86**

See notes to Table 14.21

Table 14.23: Two-factor EFA for Whites and Indians on the child SDQ

	White (N=6872)		**Indian (N=184)**	
Item	**“Internalising” factor**	**“Externalising” factor**	**“Internalising” factor**	**“Externalising” factor**
Somatic [i]	**-0.43**		**0.41**	
Worries [i]	**-0.69**		**0.60**	
Unhappy [i]	**-0.70**		**0.58**	
Clingy [i]	**-0.50**		**0.46**	
Fears [i]	**-0.68**		**0.66**	
Solitary [i]	**-0.49**		0.37	
Good friend [i] *				
Popular [i] *			-0.31	
Bullied [i]	**-0.61**		**0.82**	
Best with adults [i]	**-0.38**			
Tempers [e]		**0.48**		**0.43**
Obedient [e] *		**-0.63**		**-0.64**
Fights [e]		**0.55**		0.37
Lies [e]		**0.54**	0.36	**0.41**
Steals [e]		**0.47**	**0.70**	
Restless [e]		**0.56**		**0.57**
Fidgety [e]		**0.59**		**0.63**
Distractible [e]		**0.65**	0.31	**0.53**
Reflective [e] *		**-0.62**		**-0.65**
Persistent [e] *		**-0.68**		**-0.50**

See notes to Table 14.21

14.5.4 MPlus syntax for multi-group confirmatory factor analysis for 20-item two-factor general-specific model

```
TITLE: Multi-group CFA for parent SDQ of 20 item two-factor general-
specific model

DATA:
    FILE IS [name_of_datafile];

VARIABLE:
    NAMES ARE studyno weight strata cluster indian
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25;

    USEVARIABLES ARE
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20;

   CATEGORICAL ARE
      p1 p2 p3 p4 p5 p6 p7 p8 p9 p10
      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20;

    missing are all  (999);
    stratification is strata;
    cluster are cluster;
    weight is weight ;
    idvar is studyno;

    grouping is indian (0 = White 1 = Indian);

ANALYSIS:
    TYPE is complex missing h1;
    ESTIMATOR IS wlsmv;
    ITERATIONS = 100000;
    CONVERGENCE = 0.0005;

MODEL:
    internalising by p1*1 p2 p3 p4 p5 p6 p7 p8 p9 p10;
    externalising by p11*1 p12 p13 p14 p15 p16 p17 p18 p19 p20;

    all           by  p1*1 p2 p3 p4 p5 p6 p7 p8 p9 p10
                      p11 p12 p13 p14 p15 p16 p17 p18 p19 p20;

    internalising with externalising @0;
    internalising with all @0;
    externalising with all @0;

    internalising @1;
    externalising @1;
    all @1;

OUTPUT:
    SAMPSTAT STANDARDIZED;
```

Key to on variable names: studyno=unique individual ID; weight=individual probability weight; strata= strata ID; cluster=cluster ID; Indian=Indian (vs. White) ethnicity; p1-p5=five SDQ emotional problems items; p6-p10=five SDQ peer problems items; p11-p15=five SDQ behavioural problems items; p16-p10=five SDQ hyperactivity items; p21-p25=five SDQ prosocial items.

14.5.5 Comparing validity coefficients from MTMM analyses between Indians and Whites

Methods

Comparing the Spearman's coefficients between Indians and Whites is not straightforward. This is because, just as when comparing symptoms and impact/burden, the lower variance in the externalising scores of Indians means one would expect them to show lower agreement between measures [496]. This reflects the fact that positively skewed scales such as the SDQ are better at distinguishing between slightly different levels of mental health at the high 'problem' end where there is a large range of possible scores, than at the low end where most children are compressed across a very narrow range of scores.

I therefore additionally performed MTMM analyses for the subsample of the White informants who were frequency matched to the Indians for parent SDQ score.[47] I compared this validity coefficient to the Indian sample, after using the Fisher's z-transformation to transform the coefficients into values with an approximately normal distribution (see Appendix 1, p.360).

Results

Table 8.10 shows the MTMM analyses for the full White sample, the White sample frequency-matched on the parent SDQ and the full Indian sample. In all three samples, the validity coefficients (grey) are larger than other coefficients in the same heterotrait block, indicating good convergent and discriminant validity.

Frequency matching for parent SDQ did not achieve complete equality between Indians and Whites for teacher- and child- externalising scores. In teachers the mean externalising score was 3.7 in the full White sample, 3.1 in the matched White sample and 2.7 in the Indians sample. In children the corresponding values were 6.0, 5.7 and 4.9. Despite this, the absolute values of the correlation coefficients are generally similar between Indians and the frequency-matched White sample or else only slightly lower. The one exception is that

[47] I also explored using other more complicated systems of matching, including frequency matching simultaneously for more than one informant. These produced very similar substantive findings, however, and I therefore chose to present results using the more transparent method of matched on the parent SDQ alone.

in Indians child-reported externalising scores showed a substantially lower correlation with the teacher subscales. This was not replicated in relation to the parent-child heterotrait block, however, and in the context of multiple testing may represent a chance finding. Certainly the teacher and the child externalising SDQ scales ‘agree’ with each other in the sense that both indicated a substantial Indian advantage in Chapter 7. I therefore conclude that there is no convincing evidence that the absolute values of the correlation coefficients are lower in Indians than Whites.

Table 14.24: MTMM analyses for Whites and Indians for the internalising and externalising subscales

			Parent		Teacher		Child	
			Int	Ext	Int	Ext	Int	Ext
Whites(full sample; presented in main body of the thesis)	Parent	Int						
		Ext	0.37					
	Teacher	Int	0.31	0.22				
		Ext	0.14	0.49	0.37			
	Child	Int	0.41	0.18	0.26	0.08		
		Ext	0.21	0.49	0.17	0.38	0.38	
White (frequency matched on parent SDQ)	Parent	Int						
		Ext	0.36					
	Teacher	Int	0.32	0.20				
		Ext	0.10	0.41	0.35			
	Child	Int	0.43	0.20	0.27	0.07		
		Ext	0.16	0.43	0.15	0.33	0.37	
Indian (full sample; presented in main body of the thesis)	Parent	Int						
		Ext	0.40					
	Teacher	Int	0.21	0.10				
		Ext	0.14	0.30*	0.26			
	Child	Int	0.41	0.23	0.13	0.01		
		Ext	0.20	0.35	-0.06**	0.07**	0.36	

*p<0.05, **p<0.01, ***p<0.001, testing for equality between Indians and the matched White sample. Int=internalising, ext=externalising. Heterotrait blocks are outlined in bold, validity coefficients are in cells shaded grey.

14.5.6 Construct validity of the SDQ subscales relative to the DAWBA bands: non-proportional odds ratios

Table 14.25: Internalising and externalising SDQ subscales as independent predictors of emotional, behavioural and hyperactivity DAWBA bands in Indians and Whites: non-proportional OR & 95% CI

			Informant-specific emotional DAWBA band		**Informant-specific behavioural DAWBA band**		**Informant-specific hyperactivity DAWBA band**	
			Whites (N=16338)	**Indians (N=365)**	**Whites (N=12312)**	**Indians (N=293)**	**Whites (N=6781)**	**Indians (N=181)**
Parent	**Internalising SDQ**	**1 vs. 2/3/4**	1.58 (1.55, 1.61)***	1.43 (1.23, 1.67)***	0.98 (0.96, 0.99)***	0.93 (0.86, 1.01)	1.11 (1.09, 1.13)***	1.09 (0.95, 1.24)
		1/2 vs. 3/4	1.37 (1.34, 1.39)***	1.23 (1.09, 1.39)**	1.11 (1.09, 1.13)***	1.01 (0.88, 1.15)	1.15 (1.13, 1.18)***	[1.11 (0.96, 1.30)]
		1/2/3 vs. 4	1.41 (1.37, 1.45)***	[1.37 (1.18, 1.59)***]	1.17 (1.14, 1.20)***	[1.17 (1.00, 1.37)]	1.16 (1.12, 1.20)***	[Single case]
	Externalising SDQ	**1 vs. 2/3/4**	1.02 (1.01, 1.03)**	1.04 (0.92, 1.17)	1.41 (1.39, 1.43)***	1.35 (1.24, 1.46)***	1.69 (1.65, 1.72)***	1.55 (1.34, 1.79)
		1/2 vs. 3/4	1.06 (1.05, 1.08)***	1.21 (1.07, 1.37)**	1.59 (1.56, 1.63)***	1.53 (1.35, 1.74)***	1.59 (1.55, 1.63)***	[1.59 (1.27, 2.00)]
		1/2/3 vs. 4	1.09 (1.06, 1.11)***	[1.52 (1.08, 2.14)*]	1.63 (1.59, 1.68)***	[1.74 (1.46, 2.07)***]	[.67 (1.60, 1.73)***	[Single case]
Teacher	**Internalising SDQ**	**1 vs. 2/3/4**	1.48 (1.45, 1.52)***	1.60 (1.38, 1.85)***	1.10 (1.08, 1.12)***	1.01 (0.92, 1.11)	1.10 (1.07, 1.12)***	1.15 (1.01, 1.30)
		1/2 vs. 3/4	1.62 (1.55, 1.70)***	[Single case]	1.18 (1.14, 1.21)***	[1.04 (0.81, 1.32)]	1.08 (1.05, 1.11)***	0.99 (0.78, 1.24)
		1/2/3 vs. 4	–	–	1.17 (1.13, 1.21)***	[1.28 (0.88, 1.87)]	1.05 (1.01, 1.09)***	[Two cases]
	Externalising SDQ	**1 vs. 2/3/4**	1.18 (1.16, 1.20)***	1.35 (1.19, 1.53)***	1.60 (1.57, 1.63)***	1.60 (1.39, 1.84)***	1.81 (1.77, 1.86)***	1.85 (1.53, 2.24)**
		1/2 vs. 3/4	1.13 (1.08, 1.18)***	[Single case]	1.63 (1.59, 1.67)***	[1.71 (1.40, 2.09)***]	1.84 (1.79, 1.90)***	2.11 (1.68, 2.66)**
		1/2/3 vs. 4	–	–	1.67 (1.62, 1.72)***	[1.79 (1.33, 2.40)***]	1.73 (1.66, 1.81)***	[Two cases]
Child	**Internalising SDQ**	**1 vs. 2/3/4**	1.45 (1.41, 1.48)***	1.49 (1.29, 1.72)***	0.99 (0.97, 1.01)	1.11 (0.98, 1.26)	–	–
		1/2 vs. 3/4	1.32 (1.28, 1.35)***	1.51 (1.30, 1.76)***	0.96 (0.93, 1.00)*	[0.99 (0.74, 1.34)]	–	–
		1/2/3 vs. 4	1.40 (1.34, 1.47)***	[1.27 (0.95, 1.69)]	0.91 (0.84, 0.98)**	[0.96 (0.74, 1.26)]	–	–
	Externalising SDQ	**1 vs. 2/3/4**	1.09 (1.07, 1.11)***	1.14 (0.99, 1.31)	1.35 (1.32, 1.38)***	1.19 (1.06, 1.33)**	–	–
		1/2 vs. 3/4	1.14 (1.11, 1.17)***	1.20 (0.99, 1.45)	1.45 (1.40, 1.49)***	[1.26 (0.98, 1.61)]	–	–
		1/2/3 vs. 4	1.10 (1.06, 1.15)***	[1.34 (1.02, 1.77)*]	1.49 (1.42, 1.56)***	[1.44 (0.98, 2.12)]	–	–

*p<0.05, **p<0.01, ***p<0.001 Subscales shaded grey are expected *a priori* to be the strongest predictors. Results presented in square brackets are predicting to an outcome recorded in fewer than ten children. Note no hyperactivity DAWBA band exists for children and that the teacher emotional DAWBA band does not have a fourth level.

14.6 Appendix to Chapter 9: Fuller description of measures and constituent items

14.6.1 Informants and dataset availability for B-CAMHS variables

Table 14.26: Summary of informants and dataset availability of child, family, school and area variables in B-CAMHS

Domain	Description	Source	1999	1999 +3	2004	2004 +3
Ethnicity	Ethnicity of child	P	√	√	√	√
A priori confounders	Sex[48]	P	√	√	√	√
	Age	P	√	√	√	√
	Survey year	ONS	√	√	√	√
Area	Country and region	ONS	√	√	√	√
	Index of multiple deprivation	AG	√		√	
	Indian ethnic density	AG	√		√	
School	Ford Score	ONS/AG	√		√	
Family SEP	Parent's education	P	√		√	
	Household income	P	√		√	
	Rented housing tenure	P	√		√	
	Occupational social class	P	√		√	
	Mother's economic activity	P	√		√	
	Father's economic activity	P	√		√	
Family composition and stress	Family type	P	√		√	
	Parent marital status	P	√		√	
	Three generation family	P	√		√	
	Number of co-resident siblings	P	√		√	
	Mother's age at child's birth	P	√		√	
	Parent's mental health	P (laptop)	√	√	√	√
	Family functioning	P (laptop)	√	√	√	
	Stressful life events which affect the family	P	√	√	√	√
Child	Health					
	General health	P	√	√	√	√
	Specific health complaints	P	√	√	√	
	Stressful life events specific to the child	P	√	√	√	√
	Substance use					
	Smoking	C (laptop)	√	√	√	√
	Alcohol	C (laptop)	√	√	√	√
	Drug use	C (laptop)	√	√	√	√
	Academic abilities					
	Learning difficulties	P	√	√	√	
	Dyslexia	P	√	√	√	
	Formal assessments of academic ability	I	√			
	Parent assessment of academic ability	P		√	√	√
	Teacher assessment of academic abilities	T	√	√	√	√
	Internalising mental health problems	P, T, C	√	√	√	√

[48] In a few cases (N<100) the child's sex as collected by ONS differed from that in the CBR records. In such cases, I inspected the child's first name to determine their sex.

	Parenting strategies					
	Rewards	P	√	√		
	Punishments	P	√	√		
	Relationships with peers					
	Parent’s opinion of friends	P			√	√
	Social aptitudes scale	P			√	√
	Relationships with relatives					
	Social support	C			√	(√)
	Number of close relatives	C			√	(√)
	Helping relatives	C			√	(√)

P=parent-reported, verbal interview; P (laptop)=parent-reported, self-completed on laptop; T=teacher-reported, postal questionnaire; C=child-reported, C (laptop)=child-reported; self-completed on laptop; I=assessed by interviewer; ONS=calculated or held by ONS and/or the original B-CAMHS team; AG=created by the PhD candidate. (√) indicates variables which were collected, but which I was not able to access from the B-CAMHS team.

14.6.2 Supplementary descriptions of child, family, school and area characteristics

<u>Geographical region and Metropolitan area</u>

Table 14.27: Grouping of Government Office Regions with Metropolitan counties

Government Office Region, with Metropolitan areas	Geographic region	Metropolitan vs. non-metropolitan
North East Met	North East	Metropolitan
North East Non Met	North East	Non-metropolitan
North West Met	North West	Metropolitan
North West Non Met	North West	Non-metropolitan
Merseyside	North West	Metropolitan
York and Humberside Met	York and Humberside	Metropolitan
York and Humberside Non Met	York and Humberside	Non-metropolitan
East Midlands	East Midlands	Non-metropolitan
West Midlands Met	West Midlands	Metropolitan
West Midlands Non Met	West Midlands	Non-metropolitan
Eastern Outer Met	East Anglia	Metropolitan
Eastern Other	East Anglia	Non-metropolitan
London Inner	London	Metropolitan
London Outer	London	Metropolitan
South East Outer Met	South East	Metropolitan
South East Other	South East	Non-metropolitan
South West	South West	Non-metropolitan
Wales 1 – Glamorgan, Gwent	Wales	Non-metropolitan
Wales 2 – Clwydd, Gwenneyd, Dyfed, Powys	Wales	Non-metropolitan
Scotland 1 – Highlands, Grampian, Tayside	Scotland	Non-metropolitan
Scotland 2 – Fife, Central Lothian	Scotland	Non-metropolitan
Scotland 3 – Glasgow Met	Scotland	Metropolitan
Scotland 3 – Strathclyde Exc. Glasgow	Scotland	Non-metropolitan
Scotland 4 – Borders, Dumfries, Galloway	Scotland	Non-metropolitan

English Indices of Multiple Deprivation

Table 14.28: Domains and indicators of the English Indices of Multiple Deprivation, 2004

Domain	Indicator variables
Income Deprivation	1. Adults and children in Income Support households (2001). 2. Adults and children in Income Based Job Seekers Allowance households (2001). 3. Adults and children in Working Families Tax Credit households whose equivalised income (excluding housing benefits) is below 60% of median before housing costs (2001). 4. Adults and children in Disabled Person's Tax Credit households whose equivalised income (excluding housing benefits) below 60% of median before housing costs (2001). 5. National Asylum Support Service supported asylum seekers in England in receipt of subsistence only and accommodation support (2002).
Employment Deprivation	6. Unemployment claimant count (JUVOS) of women aged 18-59 and men aged 18-64 averaged over 4 quarters (2001). 7. Incapacity Benefit claimants (women aged 18-59 and men aged 18-64) (2001). 8. Severe Disablement Allowance claimants (women aged 18-59, men aged 18-64) (2001). 9. Participants in New Deal for 18-24 year olds not included in the claimant count (2001). 10. Participants in New Deal for 25 year olds not included in the claimant count (2001). 11. Participants in New Deal for Lone Parents aged 18 and over (2001).
Health Deprivation and Disability	12. Years of Potential Life Lost (1997-2001). 13. Comparative Illness and Disability Ratio (2001). 14. Measures of emergency admissions to hospital (1999-2002). 15. Adults under 60 suffering from mood or anxiety disorders (1997-2002).
Education, Skills and Training Deprivation	16. Average points score of children at Key Stage 2 (2002). 17. Average points score of children at Key Stage 3 (2002). 18. Average points score of children at Key Stage 4 (2002). 19. Proportion of young people not staying on in school or school level education above 16 (2001). 20. Proportion of those aged under 21 not entering Higher Education (1999-2002). 21. Secondary school absence rate (2001-2002). 22. Proportions of working age adults (aged 25-54) in the area with no or low qualifications (2001).
Barriers to Housing and Services	23. Household overcrowding (2001). 24. LA level percentage of households for whom a decision on their application for assistance under the homeless provisions of housing legislation has been made, assigned to SOAs (2002). 25. Difficulty of Access to owner-occupation (2002). 26. Road distance to GP premises (2003). 27. Road distance to a supermarket or convenience store (2002). 28. Road distance to a primary school (2001-2002). 29. Road distance to a Post Office (2003).
Crime	30. Burglary (4 recorded crime offence types, April 2002-March 2003). 31. Theft (5 recorded crime offence types, April 2002-March 2003, constrained to CDRP level). 32. Criminal damage (10 recorded crime offence types, April 2002-March 2003). 33. Violence (14 recorded crime offence types, April 2002-March 2003).
The Living Environment	34. Social and private housing in poor condition (2001). 35. Houses without central heating (2001). 36. Air quality (2001). 37. Road traffic accidents involving injury to pedestrians and cyclists (2000-2002).

Source: Office of the Deputy Prime Minister, 2004 [482]

Justification of choice of IMD over other measures of area deprivation

I chose to use overall IMD as a measure of area deprivation because I believe it has several theoretical and practical advantages over other longer-established measures such as the Jarman index [678] and the Carstairs index [679]. These advantages include having a strong theoretical underpinning, being calculated using a substantially larger number of variables, having more recently been developed and validated, and being based upon indicators than can be updated annually, rather than relying on ten-yearly census information. That the IMD is designed explicitly to measure area deprivation also makes it more suitable than ACORN, which was used in previous analyses of B-CAMHS99 [2-3, 500]. ACORN is a commercially-available system of geodemographic classification which aims to guide consumer targeting by businesses (http://www.caci.co.uk/acorn/).

A further disadvantage of the Jarman index is that it includes one ethnicity-based indicator (‘proportion of household heads born in the New Commonwealth’). It is therefore not desirable as an explanatory variable in this PhD, given that it is partially composed by a *description* of the main exposure of interest. By the same logic, the inclusion of a health domain in the IMD 2004 risks circularity when predicting to a health outcome. The four indicators of the health domain do not, however, include a child health outcome. Moreover, a previous investigation of this issue found excellent agreement between the IMD 2004 and the IMD-minus-health scores, and minimal difference between the two as predictors of health outcomes [680].

Supplementary information on the Ford score

Creation of the Ford score

The Ford score is a predictor of the prevalence of emotional and behavioural problems in mainstream schools based upon routine data collected annually by the Office for Standards in Education (OFSTED). It was developed by Ford *et al.* based upon the 7864 children who attended state-funded mainstream schools in England in B-CAMHS99 [500]. Ford *et al.* created it by assessing a range of school-level variables and retaining those which were most predictive of mental disorders in B-CAMHS99, weighting each variable according to its importance. Table 14.29 summarises the four variables which make up the final Ford score, and how these are combined to give a total score.

Table 14.29: The variables, banding and scores used to create the Ford score

Indicator	**Banding**	**Points: summed to create the Ford score**
Children eligible for free school meals as a percentage of all pupils	0 – 4.99%	0
	5.00 – 9.99%	1
	10.00 – 19.99%	2
	20.00 – 29.99%	3
	30.00 – 49.99%	4
	>50.00%	5
Percentage of all pupils with statemented special educational needs	0 – 9.99%	0
	10.00 – 19.99%	2
	20.00 – 31.99%	4
	>32.00	6
Unauthorised absence rate as a percentage of all pupils	0 – 0.25%	0
	0.26 – 1.99%	1
	>2.00%	2
Exclusion rate as a percentage of all pupils	0 – 0.79%	0
	0.8 – 0.99%	2
	>1.00%	4

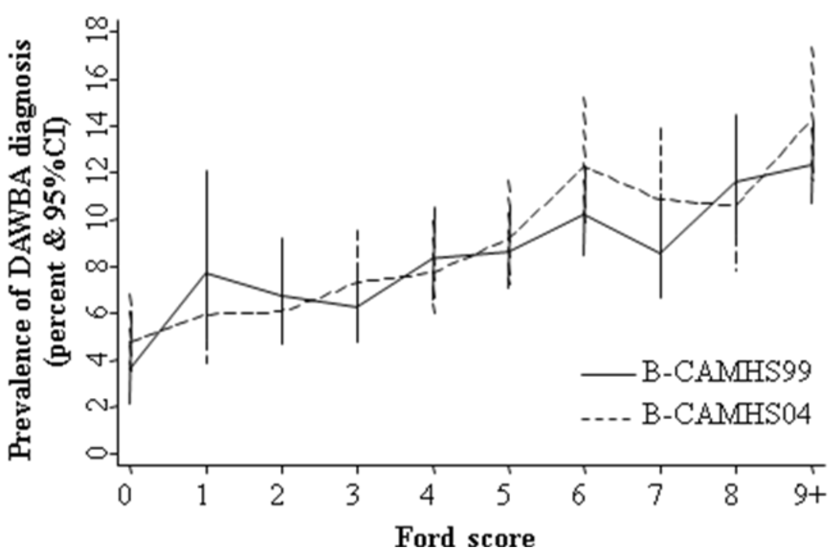
Validation of the Ford score using data from B-CAMHS04

An important limitation of validating the Ford score in B-CAMHS99 was the circularity inherent in the fact that the score was derived by identifying the factors most associated with mental disorder in those same children [417]. B-CAMHS04 therefore represented an opportunity for me conduct an independent validation of the Ford score. I calculated Ford scores for 6445 English children (91% of the English B-CAMHS sample) in collaboration with OFSTED and the Department of Children Schools and Families. To do this I linked the postcode of mainstream schools to its OFSTED Unique Reference Number (this included a small number of independent schools which underwent OFSTED inspections). I then used this to match schools to the variables listed in Table 14.29 for the academic school year 2003-4. The resulting Ford score showed a linear increase with prevalence of DAWBA diagnoses and with parent, teacher and child total difficulty SDQ scores ([501], Appendix 3). This therefore confirmed the validity of the Ford score as a predictor of the prevalence of mental health problems in schools.

Combining Ford scores between B-CAMHS99 and 04

Yet although it was again strongly associated with child mental health, the absolute value of the Ford score was lower in B-CAMHS04 (mean value 4.6) than in B-CAMHS99 (mean value 5.6, p-value for difference <0.001). I investigated whether there was any evidence that this was due to a reporting bias. If so, then one would expect any given Ford score to be associated with a higher prevalence of child mental health problems in B-CAMHS04 than in B-CAMHS99. In fact there was no evidence of such an effect. Rather, as illustrated in Figure 14.1, the prevalence of DAWBA diagnosis was very similar between B-CAMHS99 and 04 at each level of the Ford score. Logistic regression analyses confirmed that there was no evidence that survey year was associated with overall prevalence of DAWBA diagnosis (p=0.76), and this remained the case after adjusting for Ford score (p=0.26). Similar results were obtained for the parent, teacher and child TDSs. I therefore decided to combine the absolute Ford score values between B-CAMHS99 and 04.

Figure 14.1: Prevalence of DAWBA diagnosis in B-CAMHS99 and 04, stratified by Ford score



Occupational social class

Table 14.30: Operational categories of the NS-SEC, standard NS-SEC analytic classes and approximate equivalents from the Registrar General's Standard Occupational Classification (SOC) system

NS-SEC Code	Description	NS-SEC standard analytic class	Approximate SOC equivalent
L1	Employers in large organisations	1	II
L2	Higher managerial occupations	1	II
L3.1	Higher professional traditional employee	2	I
L3.2	Higher professional new employee	2	II
L3.3	Higher professional traditional self employed	2	I
L3.4	Higher professional new self employed	2	II
L4.1	Lower professional traditional employee	3	II
L4.2	Lower professional new employee	3	IIIN
L4.3	Lower professional traditional self employed	3	II
L4.4	Lower professional new self employed	3	IIIN
L5	Lower managerial occupations	3	II
L6	Higher supervisory occupations	3	IIIN
L7.1	Intermediate clerical and administrative	4	IIIN
L7.2	Intermediate sales and service	4	IIIN
L7.3	Intermediate technical and auxiliary	4	II
L7.4	Intermediate engineering	4	IIIM
L8.1	Employers in small orgs non-professional	5	II
L8.2	Employers in small orgs agriculture	5	II
L9.1	Own account workers non professional	5	IIIM
L9.2	Own account workers agriculture	5	II
L10	Lower Supervisory occupations	6	IIIM
L11.1	Lower technical craft	6	IIIM
L11.2	Lower technical process operative	6	IV
L12.1	Semi routine sales	7	IIIN
L12.2	Semi routine services	7	IV
L12.3	Semi routine technical	7	IIIM
L12.4	Semi routine operative	7	IV
L12.5	Semi routine agricultural	7	IV
L12.6	Semi routine clerical	7	IIIM
L12.7	Semi routine childcare	7	IV
L13.1	Routine sales and service	8	IV
L13.2	Routine production	8	IV
L13.3	Routine technical	8	IIIM
L13.4	Routine operative	8	V
L13.5	Routine agricultural	8	IV
L14.1	Never Worked	Never Worked	Never Worked
L14.2	Long-term unemployed	Never Worked	Never Worked
L15	Full-time students	Full-time student	Full-time student
L16	Occupation not stated or inadequately described	[missing]	[missing]
L17	Not classified for other reasons	[missing]	[missing]

Source for approximate SOC equivalents: Rose et al. [504, Appendix 2]
Designations of eight NS-SEC analytic classes: 1=Large employers and higher managerial occupations; 2=Higher professional occupations; 3=Lower managerial and professional occupations; 4=Intermediated occupations; 5=Small employers and own account workers; 6=Lower supervisory and technical occupations; 7=Semi-routine occupations; 8=Routine occupations.

Supplementary information on the cross-cultural validity of the GHQ-12

Validity and cross-cultural validity of GHQ-12

The GHQ-12 is probably the most widely used screening instrument for common mental disorders in community settings [506]. Its good reliability and validity has been demonstrated in a wide range of countries around the world [505, 507-508], with sensitivities and specificities between 73-96% when compared to clinical psychiatric interviews. Of particular relevance to this PhD has been the validation of the GHQ-12 in India [509-510] and in Indian-origin groups in the UK [511-512]. Indeed, in British Punjabis, the GHQ-12 may have greater validity than a similar measurement tool (the Amritsar Depression Inventory) originally designed in the Punjab [511]. More generally, most locally developed psychiatric screening questionnaires show a similar structure and composition to the GHQ-12 and a high agreement in case classification [681]. And while a Glasgow study found some evidence that the GHQ-12 is relatively insensitive to distress from certain sorts of stresses that disproportionately affect minority ethnic groups, it found no evidence of *differential* insensitivity for South Asians compared with Whites [682].

Psychometric properties of the GHQ-12 in the B-CAMHS survey

I investigated whether there was evidence of measurement invariance between Indians and Whites in the factor structure of the GHQ-12. Results from previous investigations have been somewhat inconsistent and provide no clear guide as to what factor structure to expect. Studies in several parts of the world report finding support for a two-factor model of 'depression' and 'social dysfunction' [e.g. 513, 514-515], but there is variation in which items load onto which factors and some studies instead find a three-factor solution [516-517].

I therefore performed an exploratory principal factor analysis for ordinal data. In both Indians and Whites, there were two factors with an Eigenvalue of greater than one and these were very similar between the two groups (see Table 14.31). The results in the pooled sample were very similar to Whites, with items 1, 3, 4, 7, 8 and 12 loading most highly on factor 1 and items 2, 5, 6, 9, 10, 11 on factor 2. This bears some resemblance to the depression (first factor)/social dysfunction (second factor) division reported in the literature, but fails to replicate previous work in several respects. For example, a WHO across 15 countries found that items 4, 5 and 6 always loaded together [513].

I then ran a series of CFAs taking these as my two factors. A general-specific model showed the best fit to the data in the pooled sample, with all items showing satisfactory loadings (>0.4) on the general factor. This model also showed adequate fit in a multigroup analysis (CFI=0.983, TLI=0.987, RMSEA=0.070), indicating measurement invariance between Indians and Whites.

Table 14.31: GHQ factor structure indicated by exploratory factor analysis in Indians and Whites

	White (N= 16 325)		**Indian (N=358)**	
GHQ-12 Items	**Factor 1**	**Factor 2**	**Factor 1**	**Factor 2**
1. Able to concentrate	0.36	0.34		0.38
2. Lost much sleep		**0.77**		**0.81**
3. Playing useful part	**0.65**		**0.68**	
4. Capable of making decisions	**0.87**		**0.85**	
5. Under stress		**0.86**		**0.92**
6. Could not overcome difficulties		**0.76**		**0.80**
7. Enjoy normal activities	**0.45**	0.36	0.33	**0.52**
8. Face up to problems	**0.64**		**0.67**	
9. Feeling unhappy and depressed		**0.84**		**0.91**
10. Losing confidence		**0.76**		**0.77**
11. Think of self as worthless		**0.73**		**0.68**
12. Feeling reasonably happy	**0.50**	0.33	**0.42**	0.34

Coefficients created after geomin rotation. Coefficients of ≥0.4 shown in bold, coefficients of <0.3 not shown.

Supplementary information on the cross-cultural validity of the GF scale

Validity and cross-cultural validity of the GF scale of the Family Activity Device

The GF scale, either with or without other subscales of the Family Activity Device, has been used to study family functioning in diverse populations and across a range of medical conditions [520]. It has been shown to have good psychometric properties in terms of its test-retest reliability and internal consistency, to have good predictive validity and to be able to distinguish well between non-clinical families and families attending a psychiatric service [519-521]. This data comes from several countries, including the UK.

The creators of the Family Activity Device describe their approach to families as “rooted in the Judaeo-Christian value system which emphasizes the optimal development of each human being” [683, p.447]. Despite this, the measure has been used to investigate child mental health with apparent success in a variety of cultural settings [520], including post-communist Hungary [522], contemporary China [523] and Hawaiian- and Japanese-Americans in the United States [524]. Nevertheless, while many of these studies provide evidence that the GF scale has discriminant validity, there has thus far been little rigorous evaluation of the GF's scale cross-cultural appropriateness and validity (with the only such work relating to the Dutch version [684]). Moreover, I know of no research investigating its psychometric properties in minority ethnic groups in the UK.

Psychometric properties of the GF scale in the B-CAMHS survey

Most previous research has focussed on investigating the factor structure of the full Family Activity Device, rather than the internal structure of just the GF scale. An exploratory principle factor analysis indicated a two-factor structure in both Indians and Whites in B-CAMHS. These two factors seemed to be tapping into valences rather than substantive constructs, with the positively worded (even number) items forming one factor and the negatively worded (odd number) items the other (see Table 14.32). A general-specific model using these two 'positive' and 'negative' factors showed the best fit to the data in the pooled sample, with all items showing satisfactory loadings (>0.4) on the general factor. In a multigroup CFA, this model showed evidence of measurement invariance between Indians and Whites (CFI=0.991, TLI=0.993, RMSEA=0.048).

Table 14.32: GF factor structure indicated by exploratory factor analysis in Indians and Whites

	White (N=16 280)		**Indian (N=346)**	
Items	Factor 1	Factor 2	Factor 1	Factor 2
1. Planning family activities is difficult because we misunderstand each other	**0.69**		**0.70**	
2. In times of crisis we can turn to each other for support		**0.59**		**0.46**
3. We cannot talk to each other about the sadness we feel	**0.77**		**0.76**	
4. Individuals are accepted for what they are		**0.71**		**0.48**
5. We avoid discussing our fears and concerns	**0.70**		**0.72**	
6. We can express feelings to each other		**0.67**		**0.82**
7. There is lots of bad feeling in the family	**0.67**		**0.77**	
8. We feel accepted for what we are		**0.77**		**0.71**
9. Making decisions is a problem for our family	**0.59**		**0.56**	
10. We are able to make decisions on how to solve problems		**0.57**		**0.69**
11. We don't get along well together	**0.67**		**0.43**	
12. We confide in each other		**0.62**		**0.67**

Coefficients of ≥0.4 shown in bold, coefficients of <0.3 not shown.

Specific health complaints

Of the health complaints listed in Table 14.33, I excluded 'hay fever' and 'other allergies' as not being associated with externalising problems (p>0.9). I also excluded 'Stomach or digestive problems or tummy pains' and 'soiling pants' as plausibly being symptoms rather than causes of externalising problems.

Table 14.33: Prevalence of specific physical health complaints used in B-CAMHS, and association with parent externalising SDQ score (N=18 180)

Condition	N (%)	Regression coefficient (95%CI) for parent externalising SDQ score (N=18 158)†
Asthma	2696 (14.8%)	0.63 (0.47, 0.79)***
Eczema	2301 (12.7%)	0.31 (0.14, 0.47)***
Hay fever	1911 (10.5%)	-0.01 (-0.20, 0.18)
Eye or sight problems	1897 (10.4%)	0.76 (0.56, 0.96)***
Some other allergy	1072 (5.9%)	-0.02 (-0.27, 0.22)
Stomach or digestive problems or tummy pains	1045 (5.7%)	1.08 (0.83, 1.33)***
Migraine or severe headaches	850 (4.7%)	1.02 (0.75, 1.28)***
Bed wetting	836 (4.6%)	2.40 (2.09, 2.72)***
Glue ear, otitis media or grommets	786 (4.3%)	0.93 (0.64, 1.21)***
Speech or language problems	697 (3.8%)	3.08 (2.73, 3.42)***
Hearing problems	661 (3.6%)	1.59 (1.25, 1.93)***
Food allergy	642 (3.5%)	0.91 (0.58, 1.24)***
Any difficulty with coordination	406 (2.2%)	3.64 (3.18, 4.10)***
Any stiffness or deformity of the foot, leg, fingers, arms or back	325 (1.8%)	1.67 (1.22, 2.12)***
A heart problem	228 (1.3%)	0.88 (0.34, 1.42)**
Soiling pants	203 (1.1%)	3.57 (2.96, 4.19)***
Kidney, urinary tract problems	201 (1.1%)	0.73 (0.15, 1.30)*
Any muscle disease or weakness	162 (0.89%)	2.13 (1.45, 2.81)***
Obesity	152 (0.84%)	1.59 (0.86, 2.31)***
Epilepsy	122 (0.67%)	3.48 (2.74, 4.23)***
A condition present since birth such as club foot or cleft palate	115 (0.63%)	1.16 (0.37, 1.96)**
Diabetes	78 (0.43%)	1.27 (0.37, 2.16)**
Any blood disorder	76 (0.42%)	0.75 (-0.31, 1.82)
Cerebral palsy	59 (0.32%)	2.28 (1.14, 3.42)***
Cancer	26 (0.14%)	0.25 (-1.32, 1.81)
Missing fingers, hands, arms, toes, feet or legs	23 (0.13%)	0.79 (-1.05, 2.62)
Cystic fibrosis	10 (0.06%)	0.91 (-1.53, 3.34)
Chronic fatigue syndrome	11 (0.06%)	-0.02 (-1.80, 1.75)
Spina bifida	6 (0.03%)	-0.50 (-2.35, 1.35)

Calculated based on the 18 180 children with complete data. Partial data was also available for a further two children. Of the 18 180 children with complete data, 18 158 had parent externalising SDQ scores. †All models adjust for child's sex, age (as a continuous variable) and survey year.

Supplementary information on measures of academic abilities

Formal assessments of cognitive/academic ability, B-CAMHS99

In B-CAMHS99, children of all ages (i.e. aged 5-15) were invited, with parental consent, to complete the following assessments of their general cognitive ability, reading ability and spelling ability:

- **General cognitive ability (IQ)** via the British Picture Vocabulary Scale, second edition (BPVS-II [526]). The BPVS-II tests a child's receptive vocabulary – the ability to recognise words when spoken. This has been found to be a strong predictor of school success and among the most important contributors to comprehensive intelligence scores (for a brief review see [526, p.2-3]). For these reasons, BPVS-II is generally considered a good screening test of scholastic aptitude and cognitive ability. Raw BPVS-II scores were age-adjusted and standardised using the standard tables, but were subsequently found to vary systematically by age. They were therefore re-standardised by the original B-CAMHS research team to achieve a mean of 100 and standard deviation of 15 across the entire age range [685].
- **Reading ability** via the Word Reading Scale of the British Ability Scales, second edition (BAS-II [527]). This tests how often children can read aloud single words printed on a card, pronouncing them correctly (with regional variations being accepted). Ability scores were generated from raw scores using the BAS-II Item Response Theory model by the original B-CAMHS team, and age-corrected standardised scores generated from BAS-II tables [2].
- **Spelling ability** via a standard BAS-II Spelling Sheet. The interviewers coded responses as correct or incorrect and software developed for the study used BAS-II decision rules to present easier or harder subsequent items, depending on how accurate previous responses were. Age-based standard ability scores were generated from raw scores in using BAS-II tables by the original B-CAMHS team, as they had been for reading ability.

Parental assessments of academic difficulties, B-CAMHS04

In B-CAMHS04 parents were asked how the reading, maths and spelling of their child compared to an average child of the same age. Response options were 'Above average' (coded 0), ‘Average’ (1), ‘Has some difficulty’ (2), ‘Marked difficulty’ (3). Parents were also asked whether their child’s school work and ability was ahead of his/her age (coded 1), about average (2), or behind his/her age (3).

Teacher assessment of academic difficulties

All teachers in B-CAMHS99 and B-CAMHS04 were asked how the reading, maths and spelling of the child compared to an average child of the same age. Response options were ‘Above average’ (coded 0), ‘Average’ (1), ‘Has some difficulty’ (2), ‘Marked difficulty’ (3). Teachers were also asked to estimate the mental age of the child in terms of intellectual and scholastic ability. I used this variable to categorise children as having an older mental age than their chronological age (40%, coded 1), the correct mental age (36%, coded 2) a mental age one to two years below their chronological age (20%, coded 3) or a mental age three years or more below their correct age (4%, coded 4). These teacher-reported variables are the only detailed measures of academic ability collected in both surveys.

Cross-cultural validity of cognitive and academic measurement variables

I could not identify any research investigating the validity of any of these measures of academic ability (including the formal tests) across ethnic groups in the UK. In my own preliminary assessment of this issue, the only measure which appeared obviously problematic was BPVS-II. This showed strong evidence of poorer performance in Indians (effect size -0.54 for Indian vs. White, $p<0.001$). By contrast, the 10 other assessments of ability either showed an Indian advantage (7 measures with evidence at $p<0.05$: formal spelling, parent spelling, parent reading, parent maths; parent overall assessment; teacher spelling and teacher maths); or a non-significant difference with the trend towards an Indian advantage (3 measures: formal reading assessment; teacher reading assessment; and teacher mental age). As Table 14.34 shows, even Indian children whose mental age was ahead of their chronological age by teacher report received BPVS-II scores well below the B-CAMHS average.

Table 14.34: Mean BPVS-II scores by teacher's assessment of mental age (B-CAMHS99 only)

	White		Indian	
	N	**Mean (95% CI)**	**N**	**Mean (95% CI)**
Mental age ahead of chronological age	4187	0.49 (0.45, 0.53)	133	-0.29 (-0.55, -0.04)
Correct mental age	4281	-0.04 (-0.07, 0.00)	83	-0.65 (-0.90, -0.39)
Mental age 1-2 years behind chronological age	2387	-0.52 (-0.57, -0.47)	56	-1.09 (-1.42, -0.76)
Mental age 3+ years behind chronological age	487	-1.12 (-1.22, -1.02)	10	-1.50 (-2.54, -0.47)

I therefore excluded the BPVS-II from subsequent analyses on the grounds of probable cross-cultural invalidity – an unexpected finding given that this is the measure the BPVS authors argues is most likely to provide a fair assessment for children with English as an additional language [526]. It is also the measure which past B-CAMHS surveys have used as a measure of academic ability [417]. By contrast, the other academic measures showed a relatively consistent picture in which Indians were advantaged for spelling and maths, but perhaps not for reading.

The Social Aptitudes Scale (SAS)

Table 14.35 presents the factor structure indicated by an exploratory factor analysis of the Social Aptitudes Scale in Indians and Whites.

Table 14.35: Factor structure of Social Aptitudes Scale indicated by exploratory factor analysis in Indians and Whites (B-CAMHS04 only)

SAS items†	White (N=6674)	Indian (N=154)
1. Able to laugh around with others, for example accepting light-hearted teasing and responding appropriately	0.63	0.70
2. Easy to chat with, even if it isn't on a topic that specially interests him/her.	0.67	0.75
3. Able to compromise and be flexible	0.76	0.75
4. Finds the right thing to say or do in order to calm a tense or embarrassing situation	0.76	0.75
5. Gracious when he/she doesn't win or get his/her own way. A good loser.	0.62	0.64
6. Other people feel at ease around him/her	0.79	0.78
7. By reading between the lines of what people say, he/she can work out what they are really thinking and feeling	0.68	0.75
8. After doing something wrong, he/she's able to say sorry and sort it out so that there are no hard feelings	0.64	0.63
9. Can take the lead without others feeling they are being bossed about	0.65	0.74
10. Aware of what is and isn't appropriate in different social situations	0.74	0.79

† Prompt: How does (Child) compare with other young people of his/her age in the following abilities? Response options: A lot worse than average (0), A bit worse than average (1), Average (2), A bit better than average (3), A lot better than average (4). A total score in the range of 0-40 is then calculated.

The social support scale

Table 14.36 presents the factor structure indicated by an exploratory factor analysis of the social support scale in Indians and Whites.

Table 14.36: Exploratory factor analyses of the seven social support items in Indians and Whites

There are people I know who…	**Full B-CAMHS sample (N= 3364)**	**White (N= 2957)**	**Indian (N=86)**
1. Make me feel loved	**0.80**	**0.80**	**0.49**
2. Make me feel happy	**0.65**	**0.64**	**0.67**
3. Accept me just as I am	**0.73**	**0.74**	**0.68**
4. Make me feel an important part of their lives	**0.82**	**0.83**	**0.81**
5. Give me support and encouragement	**0.83**	**0.84**	**0.82**
6. Would see that I am taken care of if I need to be	**0.79**	**0.79**	**0.80**
7. Can be relied on no matter what happens	**0.78**	**0.78**	**0.49**

Response options: ‘Not true’ (coded 0) ‘Partly true’ (1) or ‘Certainly true’ (2)

14.7 Appendix to Chapter 10: Supplementary data regarding preliminary substantive analyses

14.7.1 Multiple imputation model

Box 14.3: Variables included in the MICE multiple imputation mode

Components of subsequent substantive models of interest

- Outcome measures: parent and teacher externalising scores.
- All potential explanatory variables described in Section 9.2, Chapter 9. The imputation model imputed the binary variables using a logistic regression model, the pure categorical variables using multinomial logistic regression, the ordered categorical variables using ordered logistic regression and the continuous variables using linear regression.
- Quadratic and cubic terms for those continuous variables for which these were predictive of externalising problems.
- Interaction terms between Indian ethnicity and parent's education; household income; tenure; area deprivation; Indian ethnic density; family type; and three-generation family.
- Interaction terms between parent informant and parent's education; and parent's mental health.

Potential predictors of the values of the variables for which there is missing data

- Parental informant (mother, father, other).
- Other mental health measures: any DAWBA diagnosis for emotional disorder, any DAWBA diagnosis for externalising disorder, parent prosocial SDQ subscale, teacher prosocial subscale, child internalising substance, child externalising subscale and child prosocial subscale.
- All other plausible predictors of the covariates were already included in the model.

Predictors of missingness of data

- The reasons for data to be systematically missing within the English dataset were that the variable was dataset-specific, teacher-reported or logically applies only to a subset (e.g. whether the parents are married or cohabiting does not apply to lone parent families). The variables above already included all those factors found in Section 6.3, Chapter 6 to predict teacher and child non-response. These were older age in teachers, male gender in children and greater area deprivation in both informants.
- The only other variable which I knew to be systematically missing was the Ford score, which was not calculated for special needs schools and was only calculated for some independent schools. All plausible predictors of attending a special or independent school (e.g. neuro-developmental disorder, family SEP) were already included in the model.

14.7.2 Descriptive analysis of Indians and Whites in B-CAMHS

Table 14.37: Descriptive analysis of the child, family, school and area characteristics of the Whites and Indians in the sample

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
Ethnicity	**Indian ethnicity**	0%	White (n=13868)	100.0	0.0	
			Indian (n=361)	0.0	100.0	
Child mental health	**Parent-reported externalising SDQ score**	0%	Range 0-20 points	m=4.98	m=3.90	<0.001 (y)
			0-1 (2778 White, 98 Indian)	20.0	27.1	<0.001 (x)
			2-3 (3082 White, 83 Indian)	22.3	23.0	
			4-5 (2752 White, 80 Indian)	19.9	22.2	
			6-7 (1976 White, 57 Indian)	14.2	15.7	
			8-9 (1421 White, 24 Indian)	10.3	6.7	
			10-11 (838 White, 12 Indian)	6.1	3.5	
			12-13 (571 White, 4 Indian)	4.1	1.1	
			14-15 (263 White, 2 Indian)	1.9	0.5	
			16-17 (128 White, 0 Indian)	0.9	0.0	
			18-20 (59 White, 1 Indian)	0.4	0.3	
	Teacher-reported externalising SDQ score	21.7% (0.43% of interviews where teachers took part)	Range 0-20 points	m=3.70	m=2.69	<0.001 (y)
			0-1 (4154 White, 122 Indian)	38.6	47.9	<0.04 (x)
			2-3 (2428 White, 56 Indian)	22.5	21.7	
			4-5 (1473 White, 34 Indian)	13.7	13.2	
			6-7 (930 White, 22 Indian)	8.6	8.6	
			8-9 (636 White, 12 Indian)	5.9	4.5	
			10-11 (470 White, 7 Indian)	4.3	2.6	
			12-13 (323 White, 1 Indian)	3.0	0.4	
			14-15 (207 White, 2 Indian)	1.9	0.8	
			16-17 (102 White, 1 Indian)	0.9	0.4	
			18-20 (52 White, 0 Indian)	0.5	0.0	
A priori confounders	**Child's sex**	0%	Male (7056 White, 189 Indian)	50.8	52.5	0.54 (x)
			Female (6984 White, 172 Indian)	49.2	47.6	
	Child's age	0%	Range 5-16 years	m=10.2	m=10.3	0.29 (z)
			5-6 (2402 White, 55 Indian)	17.3	15.2	0.69 (x)
			7-8 (2476 White, 57 Indian)	17.8	16.3	
			9-10 (2562 White, 71 Indian)	18.2	19.3	
			11-12 (2465 White, 75 Indian)	17.7	20.4	
			13-14 (2325 White, 59 Indian)	17.2	16.7	
			15-16 (1638 White, 44 Indian)	11.9	12.2	
	Survey year	0%	1999 (7872 White, 194 Indian)	58.0	54.5	0.50 (x)
			2004 (5996 White, 167 Indian)	42.0	45.5	
Area	**Geographical region**	0%	South East (2409 White, 26 Indian)	17.4	7.0	<0.001 (x)
			London (1104 White, 109 Indian)	8.9	33.0	
			South West (1643 White, 5 Indian)	11.6	1.3	
			Eastern (1611 White, 14 Indian)	11.6	3.8	
			East Midlands (1217 White, 72 Indian)	8.5	18.6	
			West Midlands (1458 White, 61 Indian)	10.5	16.3	
			North East (788 White, 3 Indian)	5.9	0.9	
			North West & Merseyside (2155 White, 58 Indian)	15.5	16.1	
			Yorkshire & Humberside (1483 White, 13 Indian)	10.7	3.6	

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
	Metropolitan region	0%	Non-Metropolitan (7820 White, 117 Indian)	55.6	30.7	<0.001 (x)
			Metropolitan (6048 White, 244 Indian)	44.4	69.3	
	Area deprivation	0.06%	Range 0.59 – 82.3 points	m=21.0	m=26.9	<0.001 (y)
			0-10 (3901 White, 56 Indian)	28.1	15.5	<0.001 (x)
			10-20 (4407 White, 90 Indian)	31.7	25.4	
			20-30 (2310 White, 68 Indian)	16.7	18.9	
			30-40 (1379 White, 72 Indian)	10.0	19.9	
			40-50 (820 White, 38 Indian)	6.0	10.6	
			50-60 (665 White, 23 Indian)	4.8	6.2	
			60-70 (287 White, 12 Indian)	2.1	3.3	
			70+ points (92 White, 1Indian)	0.7	0.3	
	Indian ethnic density	0%	Range 0-81.8%	m=1.26	m=23.2	<0.001 (z)
			<0.01% (3911 White, 3 Indian)	27.9	0.8	<0.001 (x)
			0.01-2% (8006 White, 41 Indian)	57.6	11.1	
			2-5% (1163 White, 56 Indian)	8.6	15.8	
			5-15% (651 White, 72 Indian)	4.9	20.3	
			15-50% (133 White, 133 Indian)	1.0	37.2	
			50-100% (4 White, 56 Indian)	0.0	14.9	
School	**Ford score**	9.9%	Range 0-17 points	m=4.98	m=5.32	0.17 (y)
			0-2 (2931 White, 69 Indian)	23.3	23.0	0.08 (x)
			3-5 (4606 White, 85 Indian)	36.8	27.9	
			6-8 (3293 White, 103 Indian)	26.5	33.9	
			9-11 (1326 White, 36 Indian)	10.7	12.1	
			12-14 (294 White, 9 Indian)	2.4	2.9	
			15-17 (51 White, 1 Indian)	0.4	0.4	
Family SEP	**Parent's highest educational qualification**	0.4%	Range 'no qualifications' (coded 0) to 'degree level' (coded 6)	m=3.15	m=3.08	0.05 (z)
			No qualifications (2717 White, 102 Indian)	19.8	28.3	<0.001 (x)
			Poor GCSEs (2063 White, 64 Indian)	14.9	17.7	
			Good GCSEs (4337 White, 68 Indian)	31.3	18.9	
			A-level (1487 White, 24 Indian)	10.7	6.8	
			Diploma (1496 White, 35 Indian)	10.8	9.7	
			Degree (1715 White, 65 Indian)	12.5	18.6	
	Weekly household income	5.7%	Range '£0-99' (coded 0.5) to '£770+' (coded 8.5)	m=4.81	m=4.57	0.08 (z)
			£0-99 (506 White, 9 Indian)	3.9	2.9	<0.001 (x)
			£100-199 (1905 White, 34 Indian)	14.6	10.8	
			£200-299 (1727 White, 77 Indian)	13.1	24.7	
			£300-399 (1578 White, 44 Indian)	12.0	13.9	
			£400-499 (1464 White, 32 Indian)	11.1	10.6	
			£500-599 (1319 White, 23 Indian)	10.1	7.4	
			£600-769 (1802 White, 23 Indian)	13.7	7.6	
			£770 and over (2806 White, 67 Indian)	21.5	22.3	
	Housing tenure	0.04%	Owner occupied (9854 White, 320 Indian)	71.0	88.7	<0.001 (x)
			Social sector rented (3109 White, 27 Indian)	22.5	7.7	
			Privately rented (901 White, 13 Indian)	6.5	3.6	
	Occupational social class	2.2%	I (747 White, 31 Indian)	5.6	9.4	0.03 (x)
			II (4125 White, 102 Indian)	30.6	30.3	
			III Non-manual (2743 White, 55 Indian)	19.9	15.7	
			III Manual (2435 White, 61 Indian)	18.1	17.9	

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
			IV (2530 White, 79 Indian)	18.5	22.9	
			V (680 White, 10 Indian)	5.0	3.0	
			Never worked (189 White, 3 Indian)	1.4	0.9	
			Full-time student (125 White, 0 Indian)	0.9	0.0	
	Mother's economic activity [nested]	0.9% of families in which the mother was present	Full-time employed (3255 White, 117 Indian)	24.2	33.3	<0.001 (x)
			Part-time employed (6204 White, 101 Indian)	46.0	28.3	
			Home and family (3134 White, 112 Indian)	23.4	31.2	
			Unemployed (352 White, 7 Indian)	2.7	2.0	
			Other (503 White, 18 Indian)	3.8	5.1	
	Father's economic activity [nested]	1.0% of families in which the father was present	Full-time employed (9511 White, 266 Indian)	87.3	79.6	0.003 (x)
			Part-time employed (353 White, 21 Indian)	3.2	6.2	
			Home and family (213 White, 9 Indian)	1.9	2.7	
			Unemployed (315 White, 15 Indian)	2.9	4.7	
			Other (509 White, 24 Indian)	4.7	6.9	
Family composition	**Family type**	0.2%	Two-parent family (9052 White, 332 Indian)	65.4	92.2	<0.001 (x)
			Step family (1689 White, 4 Indian)	12.1	1.1	
			Lone parent family (3104 White, 25 Indian)	22.4	6.7	
	Parent marital status [nested]	0% of those in traditional or step families	Married (9446 White, 334 Indian)	88.0	99.5	<0.001 (x)
			Cohabiting (1295 White, 2 Indian)	12.0	0.5	
	Three generation family	0%	No grandparent in household (13608 White, 309 Indian)	98.1	85.5	<0.001 (x)
			Grandparent in household (260 White, 52 Indian)	1.9	14.5	
	Number of co-resident siblings	0%	Range 0 to '4 or more'	m=1.27	m=1.42	0.002 (z)
			0 (2652 White, 51 Indian)	19.3	14.3	0.08 (x)
			1 (6541 White, 166 Indian)	47.1	46.4	
			2 (3261 White, 95 Indian)	23.5	25.8	
			3 (1035 White, 34 Indian)	7.4	9.4	
			4 or more (379 White, 15 Indian)	2.7	4.1	
	Mother's age at child's birth	3.4%	Range '17 or less' to '40 or more'	m=27.9	m=27.8	0.41 (y)
			≤19 (712 White, 13 Indian)	5.3	3.5	0.29 (x)
			20-24 (2902 White, 85 Indian)	21.7	23.7	
			25-29 (4622 White, 137 Indian)	34.5	38.5	
			30-34 (3561 White, 85 Indian)	26.6	24.0	
			35-39 (1349 White, 33 Indian)	10.1	9.2	
			40 or more (238 White, 4 Indian)	1.8	1.1	
Family stress	**Parent mental health**	0.5%	Range 0-12 points	m=1.71	m=1.75	0.52 (z)
			0-1 (9338 White, 238 Indian)	67.6	68.0	0.55 (x)
			2-3 (1976 White, 43 Indian)	14.3	12.3	
			4-5 (989 White, 26 Indian)	7.2	7.3	
			6-7 (642 White, 22 Indian)	4.7	6.3	
			8-9 (439 White, 8 Indian)	3.2	2.5	
			10-12 (417 White, 13 Indian)	3.0	3.7	
	Family functioning	0.9%	Range 1-3.75 points	m=1.69	m=1.80	<0.001 (z)
			1.0-1.49 (4206 White, 75 Indian)	30.5	22.2	<0.001 (x)
			1.5-1.99 (5835 White, 130 Indian)	42.5	38.1	
			2.0-2.49 (3274 White, 119 Indian)	23.8	35.5	
			2.5-2.99 (388 White, 14 Indian)	2.8	4.2	
			3.0-4.0 (60 White, 0 Indian)	0.4	0.0	

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
	Parental separation	0.2%	No (9470 White, 328 Indian) Yes (4369 White, 31 Indian)	68.5 31.5	91.6 8.4	<0.001 (x)
	Family financial Crisis	0.3%	No (11753 White, 319 Indian) Yes (2080 White, 39 Indian)	84.9 15.1	89.4 10.6	0.02 (x)
	Family police Contact	0.3%	No (12981 White, 346 Indian) Yes (855 White, 11 Indian)	93.8 6.2	96.9 3.1	0.02 (x)
	Death of parent or sibling	0.2%	No (13366 White, 351 Indian) Yes (473 White, 8 Indian)	96.6 3.4	97.7 2.3	0.27 (x)
Child	**General health**	0.01%	Range 'very bad' (coded 0) 'very good' (coded 4) Bad/very bad (88 White, 4 Indian) Fair (779 White, 26 Indian) Good (3248 White, 134 Indian) Very good (9751 White, 197 Indian)	m= 3.63 0.6 5.6 23.4 70.3	m=3.45 1.2 7.3 37.4 54.1	<0.001 (z) <0.001 (x) <0.001 (w)
	Neuro-developmental disorder	0.1%	No (13 741 White, 360 Indian) Yes (125 White, 1 Indian)	99.1 0.9	99.7 0.3	0.26 (x) 0.17 (w)
	Developmental problems	0.1%	No (12,523 White, 344 Indian) Yes (1344 White, 17 Indian)	90.3 9.7	95.3 4.7	0.001 (x)
	Common physical disorder	0.1%	No (8377 White, 239 Indian) Yes (5490 White, 122 Indian)	60.4 39.6	66.4 33.6	0.03 (x)
	Rare physical disorder	0%	No (12 978 White, 349 Indian) Yes (890 White, 12 Indian)	93.6 6.4	96.7 3.3	0.03 (x)
	Serious illness leading to hospitalisation	0.2%	No (11386 White, 319 Indian) Yes (2452 White, 40 Indian)	82.2 17.8	88.7 11.3	0.002 (x)
	Death of friend	0.2%	No (12997 White, 349 Indian) Yes (840 White, 10 Indian)	93.9 6.1	97.2 2.8	0.01 (x)
	Regular smoker	5.1% (1.0% of 11-16 year olds who took part)	No (12 7999 White, 334 Indian) Yes (363 White, 4 Indian)	97.2 2.7	98.9 1.2	0.11 (x)
	Alcohol consumption	5.2 (1.1% of 11-16 year olds who took part)	Never/rare (12 126 White, 333 Indian) Moderate (803 White, 4 Indian) Frequent (229 White, 1 Indian)	92.1 6.2 1.8	98.5 1.2 0.4	<0.001 (x)
	Ever used drugs	5.2% (1.1% of 11-16 year olds who took part)	No (12 646 White, 329 Indian) Yes (509 White, 9 Indian)	96.1 3.4	97.2 2.8	0.34 (x)
	Teacher-reported difficulties in school	24.0% (3.2% of interviews where teacher took part)	Range 0-9 points 0-1 (3352 White, 91 Indian) 2-3 (3689 White, 85 Indian) 4-5 (1461 White, 36 Indian) 6-7 (1361 White, 26 Indian) 8-9 (694 White, 13 Indian)	m=3.03 31.9 34.9 13.8 12.9 6.6	m=2.71 37.1 33.4 14.5 10.0 5.0	0.05 (z) 0.28 (x)
	Learning difficulty	0.1%	No (12 680 White, 351 Indian) Yes (5490 White, 10 Indian)	91.4 8.6	97.1 2.9	<0.001 (x)
	Dyslexia	0.1%	No (13 378 White, 359 Indian) Yes (489 White, 2 Indian)	96.4 3.6	99.5 0.5	<0.001 (x)
	Parent-reported internalising SDQ score	0%	Range 0-20 points 0-1 (4687 White, 111 Indian) 2-3 (3991 White, 100 Indian) 4-5 (2356 White, 74 Indian) 6-7 (1368 White, 32 Indian) 8-9 (785 White, 19 Indian) 10-11 (353 White, 16 Indian) 12-13 (171 White, 7 Indian) 14-15 (102 White, 1 Indian)	m= 3.33 33.8 28.8 17.0 9.9 5.7 2.6 1.2 0.7	m=3.54 30.8 27.4 20.2 9.3 5.2 4.3 2.2 0.3	0.37 (y) 0.36 (x)

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
			16-17 (41 White, 1 Indian)	0.3	0.3	
			18-20 (14 White, 0 Indian)	0.1	0.0	
Child, 1999 only	**Reward: praise [1999]**	43.4% (0.1% of interviews in 1999)	Never (23 White, 1 Indian)	0.3	0.6	<0.001 (x)
			Seldom (60 White, 3 Indian)	0.8	1.4	<0.001 (w)
			Sometimes (1211 White, 69 Indian)	15.5	35.8	
			Frequently (6573 White, 119 Indian)	83.4	62.2	
	Reward: treats [1999]	43.4% (0.1% of interviews in 1999)	Never (271 White, 15 Indian)	3.5	7.9	<0.001 (x)
			Seldom (648 White, 13 Indian)	8.3	6.9	
			Sometimes (4210 White, 114 Indian)	53.6	58.9	
			Frequently (2734 White, 51 Indian)	34.7	26.3	
	Reward: favourite things [1999]	43.6% (0.4% of interviews in 1999)	Never (339 White, 6 Indian)	4.4	3.0	0.54 (x)
			Seldom (762 White, 19 Indian)	9.7	10.0	
			Sometimes (4341 White, 116 Indian)	55.4	59.8	
			Frequently (2397 White, 52 Indian)	30.5	27.1	
	Punish: send to room [1999]	43.3% (0.1% of interviews in 1999)	Never (2012 White, 85 Indian)	25.8	44.0	<0.001 (x)
			Seldom (2727 White, 45 Indian)	34.6	22.9	
			Sometimes (2443 White, 57 Indian)	30.9	29.5	
			Frequently (686 White, 7 Indian)	8.7	3.6	
	Punish: grounding [1999]	43.4% (0.1% of interviews in 1999)	Never (3440 White, 132 Indian)	43.7	68.2	<0.001 (x)
			Seldom (2080 White, 22 Indian)	26.5	11.1	
			Sometimes (1873 White, 35 Indian)	23.8	18.2	
			Frequently (474 White, 5 Indian)	6.0	2.6	
	Punish: shouting [1999]	43.3% (0.02% of interviews in 1999)	Never (405 White, 16 Indian)	5.2	8.3	<0.001 (x)
			Seldom (1436 White, 20 Indian)	18.3	10.2	
			Sometimes (3856 White, 118 Indian)	49.0	60.7	
			Frequently (2173 White, 40 Indian)	27.6	20.7	
	Punish: smacking [1999]	43.3% (0.04% of interviews in 1999)	Never (4304 White, 119 Indian)	55.0	61.6	0.001 (x)
			Seldom (2839 White, 46 Indian)	35.8	23.6	<0.001(w)
			Sometimes (690 White, 26 Indian)	8.7	13.3	
			Frequently (36 White, 3 Indian)	0.5	1.5	
	Punish: ever hit or shake [1999]	43.3% (0.1% of interviews in 1999)	Never (7669 White, 179 Indian)	97.5	92.1	<0.001 (x)
			Ever (199 White, 15 Indian)	2.6	7.9	
Child, 2004 only	**Parent disapproval of friends [2004]**	57.5% (1.8% of interviews in 2004)	Approves a lot (4987 White, 134 Indian)	84.8	80.6	0.12 (x)
			Approves a little (809 White, 31 Indian)	13.7	18.9	0.15(w)
			Does not approve (92 White, 1 Indian)	1.5	0.5	
	Parent thinks friends are trouble [2004]	57.6% (2.0% of interviews in 2004)	None are trouble (3839 White, 131 Indian)	65.4	79.3	0.002 (x)
			A few are trouble (1923 White, 33 Indian)	32.7	20.1	0.002(w)
			Many are trouble (89 White, 1 Indian)	1.5	0.6	
			All are trouble (25 White, 0 Indian)	0.4	0	
	Social aptitudes score [2004]	56.8% (0.3% of interviews in 2004)	Range 0 to 40 points	m=24.5	m=26.6	<0.001 (y)
			0-9 (114 White, 0 Indian)	1.9	0.0	<0.001 (x)
			10-14 (188 White, 1 Indian)	3.1	0.6	
			15-19 (797 White, 21 Indian)	13.3	12.5	
			20-24 (2001 White, 56 Indian)	33.4	33.0	
			25-29 (1606 White, 32 Indian)	26.8	20.4	
			30-34 (866 White, 29 Indian)	14.6	16.9	
			35-40 (407 White, 27 Indian)	6.8	16.5	
	Social support score [2004, 11-16 year olds]	81.5% (2.1% of 11-16 year olds who took part in 2004)	Range 0 to 14 points	m=12.6	m=12.5	0.32 (z)
			0-7 (68 White, 0 Indian)	2.7	0.0	0.45 (x)
			8-9 (115 White, 3 Indian)	4.5	4.4	
			10-11 (255 White, 10 Indian)	10.0	14.0	
			12-13 (689 White, 22 Indian)	27.2	32.9	
			14 (1398 White, 32 Indian)	55.6	48.7	

Domain	Variable	% missing data	Range/categories (N)	White percent or mean	Indian percent or mean	P-value for association with ethnicity
	No. close relatives in the home [2004, 11-16 year olds]	81.4% (1.8% of 11-16 year olds who took part in 2004)	None (71 White, 2 Indian)	2.8	2.7	0.77 (x)
			One (422 White, 9 Indian)	16.6	13.3	0.86 (w)
			Two or more (2037 White, 56 Indian)	80.7	84.0	
	No. close relatives outside the home [2004, 11-16 year olds]	81.4% (1.7% of 11-16 year olds who took part in 2004)	None (296 White, 14 Indian)	11.6	19.5	0.09 (x)
			One (434 White, 8 Indian)	17.1	10.9	
			Two or more (1801 White, 46 Indian)	71.3	69.6	
	How often child helps relatives [2004, 11-16 year olds]	81.5% (2.1% of 11-16 year olds who took part in 2004)	Every day (378 White, 20 Indian)	14.8	28.0	0.06 (x)
			At least once a week (1408 White, 30 Indian)	55.7	45.9	
			At least once a month (428 White, 12 Indian)	17.1	18.3	
			Less than once a month (137 White, 4 Indian)	5.5	6.1	
			Never (172 White, 1 Indian)	6.9	1.6	

Nested analyses: Mother's economic activity was only collected in households in which the mother (or mother substitute) was present; father's economic activity where the father was present; and parent marital status in families where both were present. Some other variables were only sought in one dataset or from children aged 11-16, as indicated in square brackets. † Percentages and means calculated adjusting for survey design, hence small discrepancies between these percentages and those calculated from the raw number of individuals. ††(w)=p-value calculated using a Fisher's exact chi-squared test for association, not adjusting for survey design (x)=p-value calculated using a chi-squared test for association, adjusting for survey design; (y) p-value calculated using a T-test, adjusting for survey design; (z) p-value calculated using a Wilcoxon non-parametric test, not adjusting for survey design.

14.7.3 Additional exploratory analyses of the causes of the bimodal distribution of parent education in Indian families

As described in Chapter 10 (p.257) and Table 14.37 (p.457), the mean level of parental education was similar between Indians and Whites, but in Indians there was a higher frequency of both degree-level qualifications (18.6% vs. 12.5% in White) and no qualifications (28.3% vs. 19.8%). I conducted some exploratory analyses into the characteristics of parents which might shed light onto the fact that Indians are overrepresented in both the most advantaged and most disadvantaged groups for parent education. Restricting the analyses to mother informants changed the proportion of Indians and Whites at each educational level very little (<1%). The distinctive Indian profile therefore could not be explained by the higher proportion of father respondents in the Indian sample. When I looked within the Indian sample at the characteristics of parents in different educational groups, I found that the parent's age varied very little but that parents with degrees tended to have immigrated to the UK at older ages than those with GCSEs, A-levels or diplomas (Table 14.38). The underrepresentation of Indians in the intermediate education categories may therefore to some extent reflect their arrival from the Indian subcontinent too late to have been able to take the standard British secondary school exams.

Table 14.38: Characteristics of Indian parents at different levels of education

	Parent's mean age at time of B-CAMHS interview †		Proportion of parents born in UK or arriving before age 15 (B-CAMHS99 only)	
	N	**Mean age**	**N**	**Percent**
No qualifications	95	38.6	60	33.3%
Poor GCSEs	61	37.3	39	59.0%
Good GCSEs	65	38.4	36	72.2%
A-level	20	37.1	15	60.0%
Diploma	32	37.5	11	81.8%
Degree	57	39.1	21	23.8%
P-value for heterogeneity††		0.37		<0.001

† Parents did not include fathers in B-CAMHS04, whose current age was not available to me. †† p-value for heterogeneity calculated via ANOVA for mean age at time of interview and via chi-squared test for association for proportion born in the UK or arriving before age 15.

14.7.4 Association of risk factors with teacher externalising scores

Table 14.39: Cross-sectional associations between teacher externalising scores and child, family, school and area characteristics.

Domain	Variable	Categories	Teacher externalising score			
			Full sample	p-value	White mean	Indian mean
A priori confounders	**Child's sex**	Male	4.72	<0.001 [a]	4.75	3.43
		Female	2.61		2.63	1.87
	Child's age	5-6 years	3.97	<0.001 [b]	3.99	3.02
		7-8 years	3.83		3.84	3.23
		9-10 years	3.73		3.76	2.95
		11-12 years	3.45		3.49	1.95
		13-14 years	3.53		3.55	2.55
		15-16 years	3.38		3.40	2.11
	Survey year	1999	3.70	0.41 [a]	3.72	2.69
		2004	3.63		3.66	2.69
Area	**Geographical region**	South East	3.66	0.40 [a]	3.68	1.89
		London	3.53		3.64	2.36
		South West	3.69		3.69	[1.30]
		Eastern	3.62		3.63	2.44
		East Midlands	3.45		3.46	3.21
		West Midlands	3.87		3.90	3.07
		North East	3.74		3.74	[2.00]
		North West & Merseyside	3.78		3.80	3.00
		Yorkshire & Humberside	3.65		3.67	1.90
	Metropolitan region	Non-Metropolitan	3.62	0.16 [a]	3.62	3.11
		Metropolitan	3.75		3.80	2.50
	Indian ethnic density	<0.01%	3.81	<0.001 [c]	3.81	[1.56]
		0.01-2%	3.68		3.69	2.63
		2-5%	3.57		3.64	1.94
		5-15%	3.27		3.39	1.95
		15-50%	3.03		3.11	2.94
		50-100%	3.99		[7.81]	3.76
	Area deprivation	0-10 points	3.08	<0.001 [c]	3.09	1.89
		10-20 points	3.46		3.47	2.96
		20-30 points	3.99		4.03	2.76
		30-40 points	4.08		4.14	2.75
		40-50 points	4.74		4.78	3.87
		50-60 points	4.95		5.07	1.52
		60-70 points	5.01		5.10	[2.16]
		70+ points	5.53		5.62	[0.00]
School	**Ford score**	0-2	3.05	<0.001 [b]	3.08	1.76
		3-5	3.48		3.49	2.89
		6-8	3.88		3.91	2.77
		9-11	4.58		4.61	3.41
		12-14	4.95		5.05	[0.97]
		15-17	5.22		5.17	[7.00]
Family SEP	**Parent's highest educational qualification**	No qualifications	5.03	<0.001 [a]	5.09	3.24
		Poor GCSEs	4.06		4.08	3.33
		Good GCSEs	3.49		3.51	1.95
		A-level	3.23		3.23	2.85
		Diploma	3.00		2.99	3.41
		Degree	2.61		2.63	1.90

Domain	Variable	Categories	Teacher externalising score			
			Full sample	p-value	White mean	Indian mean
	Weekly household income	£0-99	5.28	<0.001 [a]	5.35	[1.60]
		£100-199	4.99		5.01	3.89
		£200-299	4.55		4.62	2.73
		£300-399	3.85		3.87	2.91
		£400-499	3.36		3.37	2.96
		£500-599	3.05		3.03	4.15
		£600-769	2.93		2.94	1.94
		£770 and over	2.86		2.89	2.00
	Housing tenure	Owner occupied	3.13	<0.001 [a]	3.15	2.73
		Social sector rented	5.37		5.39	2.41
		Privately rented	4.26		4.29	2.16
	Occupational social class	I	2.48	<0.001 [a]	2.49	2.30
		II	3.1		3.11	2.37
		III Non-manual	3.7		3.72	2.75
		III Manual	3.79		3.80	3.35
		IV	4.35		4.38	2.94
		V	4.82		4.84	2.88
		Never worked	5.36		5.45	1.31
		Full-time student	4.18		4.18	[empty cell]
	Mother's economic activity [nested]	Full-time employed	3.44	<0.001 [a]	3.48	2.26
		Part-time employed	3.24		3.24	2.99
		Home and family	4.37		4.42	2.64
		Unemployed	4.49		4.55	[1.42]
		Other	4.65		4.68	4.09
	Father's economic activity [nested]	Full-time employed	3.20	<0.001 [a]	3.22	2.56
		Part-time employed	3.65		3.60	4.38
		Home and family	5.27		5.30	[4.60]
		Unemployed	4.95		5.15	1.27
		Other	4.61		4.71	2.32
Family composition	**Family type**	Two-parent family	3.15	<0.001 [a]	3.17	2.61
		Step family	4.56		4.56	[6.76]
		Lone parent family	4.86		4.87	3.30
	Marital status [nested]	Married	3.22	<0.001 [a]	3.24	2.64
		Cohabiting	4.43		4.43	[4.08]
	Three generation family	No grandparent in household	3.66	0.008 [a]	3.68	2.76
		Grandparent in household	4.44		4.82	2.17
	Number of co-resident siblings	0	3.80	<0.001 [a]	3.82	2.53
		1	3.48		3.50	2.76
		2	3.55		3.58	2.25
		3	4.57		4.62	3.05
		4 or more	4.98		5.02	4.12
	Mother's age at child's birth	≤19	4.96	<0.001 [c]	5.00	[2.82]
		20-24	4.51		4.56	2.84
		25-29	3.41		3.44	2.40
		30-34	3.16		3.16	3.10
		35-39	3.16		3.18	2.11
		40 or more	3.58		3.59	[2.85]
Family stress	**Parent mental health**	0-1	3.43	<0.001 [b]	3.45	2.58
		2-3	3.89		3.92	2.64
		4-5	4.22		4.26	2.21
		6-7	4.35		4.42	2.57
		8-9	4.38		4.40	[3.23]
		10-12	5.00		4.99	[5.45]

Domain	Variable	Categories	Teacher externalising score			
			Full sample	p-value	White mean	Indian mean
	Family functioning	1.0-1.49 1.5-1.99 2.0-2.49 2.5-2.99 3.0-4.0	3.21 3.53 4.21 5.91 4.59	<0.001 [c]	3.22 3.56 4.26 6.00 4.59	2.71 2.20 3.03 [2.74] [empty cell]
	Parental separation	No Yes	3.20 4.77	<0.001 [a]	3.22 4.78	2.66 2.98
	Family financial crisis	No Yes	3.60 4.06	<0.001 [a]	3.63 4.08	2.68 2.81
	Family police contact	No Yes	3.58 5.21	<0.001 [a]	3.60 5.24	2.70 [2.74]
	Death of parent or sibling	No Yes	3.63 4.79	<0.001 [a]	3.65 4.84	2.72 [1.40]
Child	**General health**	Bad/very bad Fair Good Very good	4.57 5.16 4.24 3.36	<0.001 [b]	4.77 5.22 4.29 3.38	[0.74] 3.37 2.94 2.45
	Neuro-developmental disorder	No Yes	3.66 5.36	<0.001 [a]	3.68 5.39	2.69 [3.00]
	Developmental problems	No Yes	3.49 5.37	<0.001 [a]	3.52 5.39	2.59 4.47
	Common physical disorder	No Yes	3.51 3.93	<0.001 [a]	3.53 3.95	2.62 2.81
	Rare physical disorder	No Yes	3.65 4.04	0.006 [a]	3.67 4.08	2.74 [1.13]
	Serious illness leading to hospitalisation	No Yes	3.56 4.18	<0.001 [a]	3.58 4.21	2.77 2.12
	Death of friend	No Yes	3.63 4.31	<0.001 [a]	3.65 4.33	2.69 [2.76]
	Regular smoker	No Yes	3.55 6.78	<0.001 [a]	3.57 6.78	2.73 [6.00]
	Alcohol consumption	Never/rare Moderate Frequent	3.58 3.99 4.54	<0.001 [a]	3.61 4.00 4.54	2.75 [1.68] [empty cell]
	Ever used drugs	No Yes	3.56 5.48	<0.001 [a]	3.58 5.47	2.69 6.03
	Teacher-reported academic difficulties	0-1 2-3 4-5 6-7 8-9	1.74 3.25 4.72 6.26 8.13	<0.001 [d]	1.75 3.26 4.74 6.29 8.17	1.36 2.42 4.15 4.85 5.71
	Learning difficulty	No Yes	3.37 7.03	<0.001 [a]	3.39 7.05	2.64 [4.49]
	Dyslexia	No Yes	3.61 5.45	<0.001 [a]	3.64 5.43	2.64 8.63
	Parent-reported internalising SDQ score	0-1 2-3 4-5 6-7 8-9 10-11 12-13 14-15 16-17 18-20	3.15 3.44 3.80 4.27 4.95 4.95 6.28 7.35 6.29 6.88	<0.001 [b]	3.17 3.46 3.81 4.29 5.01 5.03 6.47 7.43 6.36 6.88	2.13 2.41 3.37 3.65 2.75 3.26 [2.15] [1.00] [5.00] [empty cell]

Domain	Variable	Categories	Teacher externalising score			
			Full sample	p-value	White mean	Indian mean
Child, 1999 only	Reward: praise	Never	3.32	0.006 [a]	3.32	[empty cell]
		Seldom	4.69		4.64	[7.00]
		Sometimes	3.97		4.02	2.96
		Frequently	3.64		3.66	2.49
	Reward: treats	Never	3.59	<0.001 [a]	3.68	2.07
		Seldom	3.33		3.34	3.22
		Sometimes	3.52		3.54	2.98
		Frequently	4.08		4.11	2.17
	Reward: favourite things	Never	3.32	0.004 [a]	3.34	[1.48]
		Seldom	3.62		3.66	2.20
		Sometimes	3.59		3.60	2.88
		Frequently	3.99		4.02	2.67
	Punish: send to room	Never	3.09	<0.001 [a]	3.12	2.47
		Seldom	3.27		3.28	2.58
		Sometimes	4.17		4.20	2.82
		Frequently	5.58		5.59	[4.58]
	Punish: grounding	Never	2.90	<0.001 [a]	2.91	2.47
		Seldom	3.53		3.54	3.01
		Sometimes	4.78		4.82	2.80
		Frequently	6.49		6.49	[6.84]
	Punish: shouting	Never	3.11	<0.001 [a]	3.16	[1.08]
		Seldom	3.13		3.15	1.90
		Sometimes	3.62		3.63	3.17
		Frequently	4.34		4.38	2.19
	Punish: smacking	Never	3.34	<0.001 [a]	3.36	2.55
		Seldom	4.01		4.03	2.95
		Sometimes	4.54		4.61	2.80
		Frequently	5.42		5.60	[3.02]
	Punish: ever hit or shake	Never	3.67	0.004 [a]	3.69	2.69
		Ever	4.81		4.98	2.70
Child, 2004 only	Parent disapproval of friends	Approves a lot	3.29	<0.001 [a]	3.31	2.62
		Approves a little	5.14		5.24	2.96
		Does not approve	7.24		7.32	[2.00]
	Parent thinks friends are trouble	None are trouble	2.93	<0.001 [a]	2.94	2.49
		A few are trouble	4.68		4.70	3.46
		Many are trouble	8.59		8.59	[empty cell]
		All are trouble	7.93		7.93	[empty cell]
	Social aptitudes score	0-9	8.81	<0.001 [c]	8.81	[empty cell]
		10-14	7.34		7.36	[5.00]
		15-19	5.12		5.21	2.05
		20-24	3.81		3.82	3.39
		25-29	2.99		3.00	2.27
		30-34	2.52		2.52	2.56
		35-40	2.23		2.22	2.53
	Social support score	0-7	6.06	<0.001 [b]	6.06	[empty cell]
		8-9	5.53		5.64	[2.66]
		10-11	4.43		4.52	[2.55]
		12-13	3.37		3.35	3.81
		14	2.66		2.68	1.88
	No. close relatives in the home	None	4.93	<0.001 [a]	4.93	[empty cell]
		One	4.31		4.33	[3.40]
		Two or more	2.98		2.98	2.70

Domain	Variable	Categories	Teacher externalising score			
			Full sample	p-value	White mean	Indian mean
	No. close relatives outside the home	None	3.24	0.002 [a]	3.22	[3.68]
		One	3.85		3.87	[2.43]
		Two or more	3.08		3.09	2.58
	How often child helps relatives	Every day	3.16	0.70 [a]	3.20	[1.72]
		At least once a week	3.16		3.16	2.93
		At least once a month	3.32		3.35	[2.29]
		Less than once a month	2.90		2.80	[5.23]
		Never	3.90		3.93	[1.00]

Nested analyses: Mother's economic activity was only collected in households in which the mother (or mother substitute) was present; father's economic activity where the father was present; and marital status where both were present. †[a]=variable entered as categorical; [b] variable entered as a linear term; [c] variable entered as linear plus quadratic terms; [d] variable entered as linear, quadratic and cubic terms. Cells in square brackets are based on fewer than 10 children. I tested for interactions entering variables as categorical, linear or linear plus quadratic in accordance with how they were modelled in the full sample The only exception was Indian ethnic density, which I modelled as a linear term because the U-shaped relation in the full sample appeared to reflect offsetting linear trends in the two groups.

14.7.5 Adjusting for child, family, school and area characteristics upon the magnitude of the Indian advantage: full details of preliminary univariable analyses

Table 14.40: Effect of adjusting for each child, family, school and area characteristics upon the parent externalising score: full results (complete case analysis)

	Variable†	N White	N Indian	Unadjusted regression coefficient††	Adjusted regression coefficient	Change
***A priori* confounders**	**Child's sex [a]**	13 868	361	1.06 (0.70, 1.41)	1.08 (0.73, 1.43)	+0.02
	Child's age [b]	13 868	361	1.09 (0.73, 1.45)	1.08 (0.73, 1.43)	-0.01
	Survey year [a]	13 868	361	1.09 (0.72, 1.45)	1.08 (0.73, 1.43)	-0.01
Area	**Geographical region [a]**	13 868	361	1.08 (0.73, 1.43)	1.08 (0.72, 1.44)	0.00
	Metropolitan region [a]	13 868	361	1.08 (0.73, 1.43)	1.09 (0.73, 1.44)	+0.01
	Indian ethnic density [c]	13 868	361	1.08 (0.73, 1.43)	0.98 (0.56, 1.40)	-0.10
	Area deprivation [c]	13 861	360	1.07 (0.72, 1.42)	1.37 (1.03, 1.72)	+0.30
School	**Ford score [b]**	12 504	303	0.97 (0.61, 1.33)	1.06 (0.69, 1.42)	+0.10
Family SEP	**Parent's highest educational qualification [a]**	13 815	358	1.06 (0.71, 1.42)	1.15 (0.81, 1.49)	+0.09
	Weekly household income [a]	13 107	309	0.87 (0.50, 1.25)	0.96 (0.58, 1.34)	+0.09
	Housing tenure [a]	13 864	360	1.08 (0.73, 1.43)	0.73 (0.37, 1.09)	-0.35
	Occupational social class [a]	13 574	341	0.98 (0.63, 1.34)	0.92 (0.58, 1.27)	-0.06
	Mother's economic activity [nested] [a]	13 448	355	1.07 (0.72, 1.43)	1.19 (0.83, 1.55)	+0.12
	Father's economic activity [nested] [a]	10 901	335	0.84 (0.47, 1.21)	0.92 (0.54, 1.29)	+0.08
Family composition	**Family type [a]**	13 845	361	1.08 (0.73, 1.43)	0.65 (0.31, 0.99)	-0.43
	Marital status [nested] [a]	10 729	333	0.87 (0.50, 1.25)	0.72 (0.34, 1.10)	-0.15
	Three generation family [a]	13 868	361	1.08 (0.73, 1.43)	1.13 (0.77, 1.49)	+0.05
	No. co-resident siblings [a]	13 868	361	1.08 (0.73, 1.43)	1.12 (0.77, 1.47)	+0.04
	Mother's age at child's birth [a]	13 384	357	1.06 (0.70, 1.41)	1.06 (0.70, 1.41)	0.00
Family stress	**Parent mental health [d]**	13 801	350	1.08 (0.71, 1.44)	1.07 (0.72, 1.42)	-0.01
	Family functioning [b]	13 763	338	1.03 (0.66, 1.39)	1.29 (0.91, 1.67)	+0.26
	Parental separation [a]	13 839	359	1.08 (0.72, 1.43)	0.72 (0.38, 1.06)	-0.36
	Family financial crisis [a]	13 833	358	1.08 (0.72, 1.43)	1.05 (0.70, 1.40)	-0.03
	Family police contact [a]	13 836	357	1.07 (0.72, 1.43)	1.02 (0.66, 1.37)	-0.05
	Death of parent or sibling [a]	13 839	359	1.08 (0.72, 1.43)	1.07 (0.71, 1.42)	-0.01
Child	**Good general health [b]**	13 866	361	1.08 (0.73, 1.43)	1.31 (0.98, 1.65)	+0.23
	Neuro-developmental disorder [a]	13 866	361	1.08 (0.73, 1.43)	1.06 (0.71, 1.41)	-0.02
	Developmental problems [a]	13 867	361	1.08 (0.73, 1.43)	0.95 (0.59, 1.31)	-0.13
	Common physical disorder [a]	13 868	361	1.08 (0.73, 1.43)	1.03 (0.68, 1.38)	-0.05
	Rare physical disorder [a]	13 868	361	1.08 (0.73, 1.43)	1.04 (0.70, 1.39)	-0.04

	Variable†	N White	N Indian	Unadjusted regression coefficient††	Adjusted regression coefficient	Change
	Serious illness leading to hospitalisation [a]	13 838	359	1.08 (0.72, 1.43)	1.03 (0.68, 1.39)	-0.05
	Death of friend [a]	13 837	359	1.08 (0.72, 1.43)	1.04 (0.69, 1.39)	-0.04
	Regular smoker [a]	13 162	338	0.97 (0.61, 1.34)	0.91 (0.54, 1.29)	-0.06
	Alcohol consumption [a]	13 158	153	0.97 (0.61, 1.34)	0.93 (0.56, 1.29)	-0.04
	Ever used drugs [a]	13 155	153	0.97 (0.61, 1.34)	0.95 (0.58, 1.31)	-0.02
	Learning difficulty [a]	13 867	361	1.08 (0.73, 1.43)	0.84 (0.49, 1.20)	-0.24
	Dyslexia [a]	13 867	361	1.08 (0.73, 1.43)	1.01 (0.67, 1.36)	-0.07
	Teacher-reported academic difficulties [b]	10 557	251	0.91 (0.47, 1.36)	0.70 (0.25, 1.14)	-0.21
	Parent-reported internalising SDQ score [b]	13 868	361	1.08 (0.73, 1.43)	1.18 (0.92, 1.45)	+0.10
Child, 1999 only	**Reward: praise [a]**	7867	192	0.80 (0.28, 1.33)	1.07 (0.57, 1.57)	+0.27
	Reward: treats [a]	7863	193	0.79 (0.27, 1.32)	0.73 (0.22, 1.25)	-0.06
	Reward: favourite things [a]	7839	193	0.80 (0.27, 1.32)	0.79 (0.27, 1.31)	-0.01
	Punish: send to room [a]	7868	194	0.80 (0.28, 1.32)	0.56 (0.02, 1.10)	-0.24
	Punish: grounding [a]	7867	194	0.80 (0.27, 1.32)	0.37 (-0.18, 0.91)	-0.43
	Punish: shouting [a]	7870	194	0.80 (0.28, 1.32)	0.70 (0.15, 1.25)	-0.10
	Punish: smacking [a]	7869	194	0.80 (0.28, 1.32)	0.84 (0.29, 1.39)	+0.04
	Punish: ever hit or shake [a]	7868	194	0.80 (0.28, 1.32)	0.92 (0.37, 1.47)	+0.12
Child, 2004 only	**Parent disapproval of friends [a]**	5888	166	1.37 (0.99, 1.75)	1.46 (1.08, 1.84)	+0.09
	Parent thinks friends are trouble [a]	5876	165	1.39 (1.01, 1.77)	1.09 (0.72, 1.46)	-0.30
	Social aptitudes score [c]	5979	166	1.44 (1.05, 1.82)	0.95 (0.41, 1.48)	-0.49
	Social support score [c]	2525	67	0.95 (0.34, 1.57)	1.03 (0.44, 1.61)	+0.08
	No. close relatives in the home [a]	2530	66	0.98 (0.36, 1.59)	0.92 (0.27, 1.56)	-0.06
	No. close relatives outside the home [a]	2531	67	0.96 (0.34, 1.57)	0.92 (0.30, 1.54)	-0.04
	How often child helps relatives [a]	2523	66	0.93 (0.31, 1.55)	0.92 (0.30, 1.55)	-0.01

Nested analyses: Mother's economic activity was only collected in households in which the mother (or mother substitute) was present; father's economic activity where the father was present; and marital status in families where both were present. † [a]=variable entered as categorical; [b] variable entered as a linear term; [c] variable entered as a linear plus quadratic term; [d] variable entered as a linear, quadratic plus cubic term, according to how they were modelled when calculating the univariable association between that variable and child mental health †† Both unadjusted and adjusted models control for child's sex, age and survey year. In those models in which one of these *A priori* confounders is the variable under examination, the unadjusted model controls only for the other two variables.

Table 14.41: Effect of adjusting for each child, family, school and area characteristics upon teacher externalising score: full results (complete case analysis)

	Variable†	N White	N Indian	Unadjusted regression coefficient††	Adjusted regression coefficient	Change
A priori confounders	**Child's sex [a]**	10 775	257	1.01 (0.63, 1.38)	1.05 (0.67, 1.43)	+0.04
	Child's age [b]	10 775	257	1.05 (0.67, 1.43)	1.05 (0.67, 1.43)	0.00
	Survey year [a]	10 775	257	1.05 (0.67, 1.43)	1.05 (0.67, 1.43)	0.00
Area	**Geographical region [a]**	10 775	257	1.05 (0.67, 1.43)	1.03 (0.64, 1.42)	-0.02
	Metropolitan region [a]	10 775	257	1.05 (0.67, 1.43)	1.09 (0.70, 1.48)	+0.04
	Indian ethnic density [c]	10 775	257	1.05 (0.67, 1.43)	1.17 (0.72, 1.63)	+0.12
	Area deprivation [c]	10 768	256	1.05 (0.67, 1.43)	1.30 (0.94, 1.66)	+0.25
School	**Ford score [b]**	10 013	233	1.08 (0.71, 1.46)	1.14 (0.78, 1.51)	+0.06
Family SEP	**Parent's highest educational qualification [a]**	10 752	255	1.04 (0.66, 1.42)	1.09 (0.73, 1.45)	+0.05
	Weekly household income [a]	10 251	221	0.95 (0.52, 1.38)	1.04 (0.61, 1.47)	+0.09
	Housing tenure [a]	10 773	256	1.05 (0.67, 1.43)	0.72 (0.32, 1.11)	-0.33
	Occupational social class [a]	10 583	244	1.03 (0.64, 1.42)	0.94 (0.57, 1.31)	-0.09
	Mother's economic activity [nested] [a]	10 466	253	1.02 (0.63, 1.41)	1.15 (0.77, 1.53)	+0.13
	Father's economic activity [nested] [a]	8596	241	0.79 (0.38, 1.19)	0.90 (0.49, 1.31)	+0.11
Family composition	**Family type [a]**	10 758	257	1.05 (0.67, 1.43)	0.63 (0.24, 1.02)	-0.52
	Marital status [nested] [a]	8351	237	0.76 (0.35, 1.18)	0.63 (0.22, 1.04)	-0.13
	Three generation family [a]	10 775	257	1.05 (0.67, 1.43)	1.15 (0.75, 1.54)	+0.10
	No. co-resident siblings [a]	10 775	257	1.05 (0.67, 1.43)	1.09 (0.70, 1.47)	+0.04
	Mother's age at child's birth [a]	10 409	254	1.02 (0.63, 1.40)	1.01 (0.62, 1.40)	-0.01
Family stress	**Parent mental health [b]**	10 745	250	1.07 (0.67, 1.46)	1.08 (0.70, 1.46)	+0.01
	Family functioning [b]	10 719	239	1.09 (0.69, 1.49)	1.22 (0.82, 1.61)	+0.13
	Parental separation [a]	10 766	256	1.05 (0.67, 1.43)	0.69 (0.30, 1.09)	-0.36
	Family financial crisis [a]	10 761	256	1.05 (0.67, 1.43)	1.03 (0.65, 1.41)	-0.02
	Family police contact [a]	10 765	255	1.03 (0.65, 1.42)	0.99 (0.60, 1.37)	-0.04
	Death of parent or sibling [a]	10 766	256	1.05 (0.67, 1.43)	1.03 (0.65, 1.41)	-0.02
Child	**General health [b]**	10 773	257	1.05 (0.67, 1.43)	1.22 (0.85, 1.59)	+0.17
	Neuro-developmental disorder [a]	10 774	257	1.05 (0.67, 1.43)	1.04 (0.66, 1.42)	-0.01
	Developmental problems [a]	10 775	257	1.05 (0.67, 1.43)	0.98 (0.60, 1.37)	-0.07
	Common physical disorder [a]	10 775	257	1.05 (0.67, 1.43)	1.03 (0.65, 1.41)	-0.02
	Rare physical disorder [a]	10 775	257	1.05 (0.67, 1.43)	1.04 (0.66, 1.41)	-0.01
	Serious illness leading to hospitalisation [a]	10 765	256	1.05 (0.67, 1.43)	1.02 (0.64, 1.40)	-0.03
	Death of friend [a]	10 763	256	1.04 (0.66, 1.43)	1.02 (0.64, 1.41)	-0.02
	Regular smoker [a]	10 368	249	0.94 (0.56, 1.32)	0.86 (0.48, 1.24)	-0.08

	Variable†	N White	N Indian	Unadjusted regression coefficient††	Adjusted regression coefficient	Change
	Alcohol consumption [a]	10 365	249	0.94 (0.56, 1.32)	0.88 (0.50, 1.26)	-0.06
	Ever used drugs [a]	10 362	249	0.94 (0.55, 1.32)	0.89 (0.51, 1.27)	-0.05
	Teacher-reported academic difficulties [d]	10 481	246	1.08 (0.68, 1.47)	0.79 (0.44, 1.14)	-0.29
	Learning difficulty [a]	10 774	257	1.05 (0.67, 1.43)	0.86 (0.49, 1.24)	-0.19
	Dyslexia [a]	10 774	257	1.05 (0.67, 1.43)	1.01 (0.63, 1.38)	-0.04
	Parent-reported internalising SDQ score [b]	10 775	257	1.05 (0.67, 1.43)	1.13 (0.78, 1.49)	+0.08
Child, 1999 only	**Reward: praise [a]**	6297	143	1.05 (0.49, 1.61)	1.14 (0.59, 1.70)	+0.09
	Reward: treats [a]	6295	144	1.04 (0.49, 1.59)	1.01 (0.46, 1.56)	-0.03
	Reward: favourite things [a]	6275	144	1.04 (0.49, 1.59)	1.04 (0.50, 1.58)	0.00
	Punish: send to room [a]	6297	145	1.06 (0.52, 1.61)	0.96 (0.39, 1.53)	-0.10
	Punish: grounding [a]	6297	145	1.07 (0.52, 1.61)	0.79 (0.20, 1.39)	-0.28
	Punish: shouting [a]	6298	145	1.07 (0.52, 1.61)	1.06 (0.50, 1.63)	-0.01
	Punish: smacking [a]	6298	145	1.07 (0.52, 1.61)	1.07 (0.52, 1.62)	0.00
	Punish: ever hit or shake [a]	6298	145	1.07 (0.52, 1.61)	1.13 (0.58, 1.69)	+0.06
Child, 2004 only	**Parent disapproval of friends [a]**	4405	112	0.98 (0.46, 1.50)	1.11 (0.60, 1.62)	+0.13
	Parent thinks friends are trouble [a]	4397	111	0.99 (0.46, 1.51)	0.73 (0.22, 1.24)	-0.26
	Social aptitudes score [c]	4472	111	1.03 (0.50, 1.55)	0.78 (0.21, 1.35)	-0.25
	Social support score [b]	1813	43	0.62 (-0.16, 1.40)	0.77 (-0.01, 1.56)	+0.15
	No. close relatives in the home [a]	1817	42	0.54 (-0.31, 1.39)	0.46 (-0.39, 1.30)	-0.08
	No. close relatives outside the home [a]	1818	43	0.63 (-0.15, 1.41)	0.59 (-0.21, 1.38)	-0.04
	How often child helps relatives [a]	1810	42	0.59 (-0.22, 1.39)	0.55 (-0.26, 1.36)	-0.04

Nested analyses: Mother's economic activity was only collected in households in which the mother (or mother substitute) was present; father's economic activity where the father was present; and marital status in families where both were present. † [a]=variable entered as categorical; [b] variable entered as a linear term; [c] variable entered as a linear plus quadratic term; [d] variable entered as a linear, quadratic plus cubic term, according to how they were modelled when calculating the univariable association between that variable and child mental health†† Both unadjusted and adjusted models control for child's sex, age and survey year. In those models in which one of these *a priori* confounders was the variable under examination, the unadjusted model controls only for the other two variables.

14.7.6 Full models underlying selected preliminary multivariable analyses

Table 14.42: Level 1 variables predicting to parent-reported externalising problems in preliminary multivariable analyses (N=14 229; 13 868 White, 361 Indian)

Domain	Variable	Categories	OR (95% CI)
Ethnicity	**Ethnicity**	Indian	1
		White	0.97 (0.55, 1.38)***
A priori confounders	**Child's sex**	Male	1
		Female	-1.35 (-1.47, -1.24)***
	Child's age	Change per year	-0.10 (-0.12, -0.08)***
	Survey year	1999	1
		2004	-0.08 (-0.24, 0.07)
Area	**Geographical region**	South East	1
		London	-0.16 (-0.44, 0.12)
		South West	-0.16 (-0.43, 0.11)
		Eastern	-0.23 (-0.48, 0.02)
		East Midlands	-0.09 (-0.39, 0.21)
		West Midlands	-0.07 (-0.37, 0.23)
		North East	-0.11 (-0.52, 0.30)
		North West & Merseyside	-0.16 (-0.40, 0.08)
		Yorkshire & Humberside	-0.09 (-0.36, 0.18)
	Metropolitan region	Non-Metropolitan	1
		Metropolitan	-0.05 (-0.21, 0.12)
	Area deprivation	Change per standard deviation	0.03 (-0.06, 0.13)
	Indian ethnic density	Change per 10% increase	-0.01 (-0.18, 0.17)
School	**Ford score**	Change per point	0.09 (0.07, 0.12)***
Family SEP	**Parent's highest educational qualification**	No qualifications	1***
		Poor GCSEs	-0.21 (-0.42, 0.01)
		Good GCSEs	-0.70 (-0.91, -0.50)
		A-level	-0.89 (-1.14, -0.64)
		Diploma	-0.97 (-1.22, -0.72)
		Degree	-1.40 (-1.66, -1.13)
	Weekly household income	£0-99	1**
		£100-199	0.16 (-0.27, 0.59)
		£200-299	0.12 (-0.31, 0.56)
		£300-399	0.08 (-0.35, 0.52)
		£400-499	0.02 (-0.43, 0.47)
		£500-599	-0.22 (-0.66, 0.22)
		£600-769	-0.30 (-0.73, 0.13)
		£770 and over	-0.32 (-0.74, 0.10)
	Housing tenure	Owner occupied	1***
		Social sector rented	1.10 (0.89, 1.31)
		Privately rented	0.67 (0.38, 0.95)
	Occupational social class	I	1***
		II	0.04 (-0.23, 0.31)
		III Non-manual	0.30 (0.01, 0.60)
		III Manual	0.23 (-0.07, 0.53)
		IV	0.39 (0.08, 0.70)
		V	0.52 (0.11, 0.93)
		Never worked	1.14 (0.54, 1.73)
		Full-time student	0.40 (-0.31, 1.10)

Table 14.43: Level 2 variables predicting to parent-reported externalising problems in preliminary multivariable analyses (N=14 229; 13 868 White, 361 Indian)

Domain	Variable	Categories	OR (95% CI)
Ethnicity	**Ethnicity**	Indian	1***
		White	1.09 (0.74, 1.45)
A priori confounders	**Child's sex**	Male	1***
		Female	-1.33 (-1.44, -1.21)
	Child's age	Change per year	-0.15 (-0.17, -0.13)***
	Survey year	1999	1*
		2004	-0.16 (-0.30, -0.02)
Family composition	**Family type**	Two-parent family	1***
		Step family	0.81 (0.56, 1.06)
		Lone parent family	0.53 (0.28, 0.77)
	Three generation family	No grandparent in household	1
		Grandparent in household	0.19 (-0.21, 0.59)
	Number of resident siblings	0	1***
		1	-0.01 (-0.18, 0.15)
		2	-0.03 (-0.22, 0.15)
		3	0.55 (0.26, 0.83)
		4 or more	0.66 (0.21, 1.11)
	Mother's age at child's birth	Change per decade	-0.96 (-1.09, -0.83)***
Family stress	**Parent mental health**	Change per point	0.13 (0.11, 0.16)***
	Family functioning	Change per standard deviation	0.84 (0.77, 0.91)***
	Parental separation	No	1**
		Yes	0.32 (0.08, 0.55)
	Family financial crisis	No	1
		Yes	0.10 (-0.10, 0.29)
	Family police contact	No	1***
		Yes	0.78 (0.48, 1.07)
	Death of parent or sibling	No	1*
		Yes	0.36 (0.01, 0.71)

Table 14.44: Level 3 variables predicting to parent-reported externalising problems in preliminary multivariable analyses (N=14 229; 13 868 White, 361 Indian)

Domain	Variable	Categories	OR (95% CI)
Ethnicity	**Ethnicity**	Indian White	1*** 0.89 (0.58, 1.19)
A priori confounders	**Child's sex**	Male Female	1*** -1.00 (-1.12, -0.89)
	Child's age	Change per year	-0.12 (-0.14, -0.11)***
	Survey year	1999 2004	1** -0.17 (-0.29, -0.04)
Child characteristics	**Good general health**	Change per point	-0.19 (-0.30, -0.07)**
	Neuro-developmental disorder	No Yes	1 -0.04 (-0.66, 0.59)
	Developmental problems	No Yes	1*** 0.55 (0.33, 0.78)
	Common physical disorder	No Yes	1* 0.13 (0.02, 0.24)
	Rare physical disorder	No Yes	1 -0.10 (-0.32, 0.13)
	Child hospitalisation	No Yes	1 0.09 (-0.07, 0.24)
	Death of friend	No Yes	1* 0.32 (0.07, 0.57)
	Regular smoker	No Yes	1*** 2.12 (1.70, 2.54)
	Alcohol consumption	Never/rare Moderate Frequent	1* 0.37 (0.13, 0.61) 0.42 (-0.05, 0.89)
	Ever used drugs	No Yes	1** 0.60 (0.24, 0.96)
	Teacher-reported academic difficulties	Change per point	0.40 (0.37, 0.42)***
	Learning difficulty	No Yes	1*** 0.95 (0.70, 1.20)
	Dyslexia	No Yes	1 -0.04 (-0.37, 0.30)
	Internalising mental health problems	Change per point	0.38 (0.36, 0.40)

14.8 Appendix to Chapter 11: Supplementary data regarding longitudinal analyses of causal directions and final multivariable analyses

14.8.1 Relationship between baseline and follow- up scores for teacher externalising mental health and putative risk factors

Table 14.45: Distribution of follow-up scores by scores at baseline

Teacher externalising score (SDQ points)		Follow-up										
		0-1	**2-3**	**4-5**	**6-7**	**8-9**	**10-11**	**12-13**	**14-15**	**16-17**	**18-20**	**Total**
Base line	**0-1**	1035	315	161	50	24	13	4	3	2	0	1607
	2-3	355	192	129	68	27	18	9	6	1	2	807
	4-5	157	98	90	52	41	25	16	7	1	0	487
	6-7	53	72	64	34	30	17	9	5	3	2	289
	8-9	26	40	33	38	23	11	13	8	7	1	200
	10-11	6	22	19	28	19	15	14	10	4	2	139
	12-13	4	13	13	12	12	16	13	5	7	3	98
	14-15	4	2	5	7	6	15	8	9	4	3	63
	16-17	0	1	2	4	8	4	4	4	3	1	31
	18-20	0	1	3	2	3	1	2	3	3	0	18
	Total	1640	756	519	295	193	135	92	60	35	14	3739

Spearman's correlation coefficient 0.53

Parent separation		Follow-up		
		No	**Yes**	**Total**
Baseline	**No**	4956	354	5310
	Yes	1700	251	1951
	Total	6656	605	7261

Spearman's correlation coefficient 0.10

Financial crisis		Follow-up		
		No	**Yes**	**Total**
Baseline	**No**	6068	208	6276
	Yes	890	97	987
	Total	6958	305	7263

Spearman's correlation coefficient 0.11

Family police contact		Follow-up		
		No	Yes	Total
Baseline	No	6692	196	6888
	Yes	344	30	374
	Total	7036	226	7262

Spearman's correlation coefficient 0.07

Death of parent or sibling		Follow-up		
		No	Yes	Total
Baseline	No	6675	402	7077
	Yes	168	19	187
	Total	6843	421	7264

Spearman's correlation coefficient 0.03

Parent mental health (GHQ points)		Follow-up						
		0-1	2-3	4-5	6-7	8-9	10-12	Total
Baseline	0-1	3795	536	238	157	89	95	4910
	2-3	598	165	121	65	41	43	1033
	4-5	233	94	68	52	37	42	526
	6-7	124	58	44	39	32	29	326
	8-9	82	39	40	24	23	36	244
	10-12	61	24	24	18	24	45	196
	Total	4893	916	535	355	246	290	7235

Spearman's correlation coefficient 0.37

Family functioning (GF scale)		Follow-up					
		1-1.49	1.5-1.99	2-2.49	2.5-2.99	3-3.99	Total
Baseline	1-1.49	355	254	42	5	2	658
	1.5-1.99	223	541	248	17	2	1031
	2-2.49	56	194	277	34	9	570
	2.5-2.99	3	9	35	24	3	74
	3-3.99	1	1	1	4	3	10
	Total	638	999	603	84	19	2343

Spearman's correlation coefficient 0.55

General health		Follow-up					
		Very bad	Bad	Fair	Good	Very good	Total
Baseline	Very bad	1	1	1	1	0	4
	Bad	0	4	20	14	3	41
	Fair	1	15	109	181	101	407
	Good	1	13	141	736	804	1695
	Very good	3	20	118	854	4138	5133
	Total	6	53	389	1786	5046	7280

Spearman's correlation coefficient 0.40

Neuro-developmental disorder		Follow-up		
		No	Yes	Total
Baseline	No	2334	7	2341
	Yes	5	17	22
	Total	2339	24	2363

Spearman's correlation coefficient 0.74

Developmental problem		Follow-up		
		No	Yes	Total
Baseline	No	2050	49	2099
	Yes	143	121	264
	Total	2193	170	2363

Spearman's correlation coefficient 0.53

Common physical disorder		Follow-up		
		No	Yes	Total
Baseline	No	1157	251	1408
	Yes	301	654	955
	Total	1458	905	2363

Spearman's correlation coefficient 0.51

Rare physical disorder		Follow-up		
		No	Yes	Total
Baseline	No	2104	101	2205
	Yes	80	78	158
	Total	2184	179	2363

Spearman's correlation coefficient 0.42

Serious illness leading to hospitalisation		Follow-up		
		No	Yes	Total
Baseline	No	5843	257	6100
	Yes	1084	79	1163
	Total	6927	336	7263

Spearman's correlation coefficient 0.05

Death of friend		Follow-up		
		No	Yes	Total
Baseline	No	6394	408	6802
	Yes	405	56	461
	Total	6799	464	7263

Spearman's correlation coefficient 0.06

Regular smoker		Follow-up		
		No	Yes	Total
Baseline	No	5813	338	6151
	Yes	23	110	133
	Total	5836	448	6284

Spearman's correlation coefficient 0.43

Alcohol consumption		Follow-up			
		Never/ rarely	Moderate	Frequent	Total
Baseline	Never/rarely	4748	811	301	5860
	Moderate	47	137	135	319
	Frequent	12	20	70	102
	Total	4807	968	506	6281

Spearman's correlation coefficient 0.42

Ever used drugs		Follow-up		
		No	Yes	Total
Baseline	No	5488	588	6076
	Yes	33	170	203
	Total	5521	758	6279

Spearman's correlation coefficient 0.40

Teacher-reported academic difficulties		Follow-up					
		0-1	2-3	4-5	6-7	8-9	Total
Baseline	0-1	640	239	25	5	1	910
	2-3	330	453	117	46	5	951
	4-5	37	157	89	55	2	340
	6-7	15	80	61	92	21	269
	8-9	1	11	14	39	61	126
	Total	1023	940	306	237	90	2596

Spearman's correlation coefficient 0.65

Learning difficulty		Follow-up		
		No	Yes	Total
Baseline	No	2049	68	2117
	Yes	92	154	246
	Total	2141	222	2363

Spearman's correlation coefficient 0.62

Dyslexia		Follow-up		
		No	Yes	Total
Baseline	No	2219	56	2275
	Yes	22	66	88
	Total	2241	122	2363

Spearman's correlation coefficient 0.62

Parent-reported internalising score (SDQ points)		Follow-up										
		0-1	2-3	4-5	6-7	8-9	10-11	12-13	14-15	16-17	18-20	Total
Base line	0-1	1499	678	216	84	26	4	6	0	0	0	2513
	2-3	781	746	351	141	54	30	8	4	0	0	2115
	4-5	234	423	269	149	67	27	13	5	1	0	1188
	6-7	76	160	181	144	71	43	19	7	3	1	705
	8-9	21	54	88	88	66	36	17	10	2	1	383
	10-11	6	27	37	29	25	25	30	10	8	1	198
	12-13	1	8	6	12	21	17	10	9	3	0	87
	14-15	0	1	8	9	10	14	9	5	5	2	63
	16-17	0	0	3	1	3	5	2	4	3	2	23
	18-20	0	0	0	2	0	1	1	1	0	1	6
	Total	2618	2097	1159	659	343	202	115	55	25	8	7281

Spearman's correlation coefficient 0.55

Teacher-reported internalising score (SDQ points)		Follow-up										
		0-1	2-3	4-5	6-7	8-9	10-11	12-13	14-15	16-17	18-20	Total
Base line	0-1	1138	352	159	69	29	14	6	5	1	0	1773
	2-3	483	210	119	45	38	16	7	10	3	0	931
	4-5	194	110	82	40	26	13	10	2	0	0	477
	6-7	91	56	51	25	29	12	5	4	1	0	274
	8-9	53	32	20	17	18	10	8	0	2	2	162
	10-11	10	13	16	8	11	4	3	2	0	0	67
	12-13	9	7	15	5	10	5	1	0	1	0	53
	14-15	0	1	1	2	4	5	0	1	0	0	14
	16-17	2	2	1	0	1	4	2	0	0	0	12
	18-20	0	2	1	0	1	0	1	0	0	0	5
	Total	1980	785	465	211	167	83	43	24	8	2	3768

Spearman's correlation coefficient 0.33

Rewards: praises		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	0	0	2	4	6
	Seldom	0	0	5	4	9
	Sometimes	0	3	93	131	227
	Frequently	3	8	165	1332	1508
	Total	3	11	265	1471	1750

Spearman's correlation coefficient 0.29

Rewards: treats		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	8	12	23	12	55
	Seldom	16	23	68	23	130
	Sometimes	31	74	601	255	961
	Frequently	10	19	261	310	600
	Total	65	128	953	600	1746

Spearman's correlation coefficient 0.30

Rewards: favourite things		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	20	10	24	20	74
	Seldom	15	30	83	33	161
	Sometimes	35	83	599	248	965
	Frequently	15	27	227	265	534
	Total	85	150	933	566	1734

Spearman's correlation coefficient 0.26

Punish: send to room		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	238	90	50	11	389
	Seldom	170	297	128	21	616
	Sometimes	75	191	250	55	571
	Frequently	24	34	60	54	172
	Total	507	612	488	141	1748

Spearman's correlation coefficient 0.44

Punish: grounding		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	481	200	106	10	797
	Seldom	147	199	113	19	478
	Sometimes	65	96	162	41	364
	Frequently	11	27	42	29	109
	Total	704	522	423	99	1748

Spearman's correlation coefficient 0.45

Punish: shouting		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	16	29	20	3	68
	Seldom	33	121	103	24	281
	Sometimes	46	184	470	140	840
	Frequently	14	52	234	261	561
	Total	109	386	827	428	1750

Spearman's correlation coefficient 0.42

Punish: smacking		Follow-up				
		Never	Seldom	Sometimes	Frequently	Total
Baseline	Never	679	67	10	0	756
	Seldom	452	283	37	0	772
	Sometimes	69	108	34	3	214
	Frequently	2	2	4	1	9
	Total	1202	460	85	4	1751

Spearman's correlation coefficient 0.44

Punish: ever hit/shake		Follow-up		
		No	Yes	Total
Baseline	No	1668	20	1688
	Yes	53	10	63
	Total	1721	30	1751

Spearman's correlation coefficient 0.21

Parent disapproval of friends		Follow-up			
		Approves a lot	Approves a little	Disapproves	Total
Baseline	Approves a lot	3674	417	50	4141
	Approves a little	396	152	28	576
	Disapproves	40	18	8	66
	Total	4110	587	86	4783

Spearman's correlation coefficient 0.21

Parent thinks friends are trouble		Follow-up				
		None	A few	Many	All	Total
Baseline	None	2660	563	24	16	3263
	A few	826	562	38	8	1434
	Many	20	28	6	3	57
	All	14	2	2	0	18
	Total	3520	1155	70	27	4772

Spearman's correlation coefficient 0.26

Social aptitudes score (SAS points)		Follow-up							
		0-9	10-14	15-19	20-24	25-29	30-34	35-40	Total
Baseline	0-9	43	21	8	2	3	1	0	78
	10-14	14	29	39	28	8	4	0	122
	15-19	14	45	163	284	89	26	10	631
	20-24	3	31	177	685	546	170	62	1674
	25-29	2	6	51	301	509	360	117	1346
	30-34	3	3	12	66	190	263	161	698
	35-40	0	0	0	15	53	113	180	361
	Total	79	135	450	1381	1398	937	530	4910

Spearman's correlation coefficient 0.62

14.8.2 Results of regression models to assess the direction of causality

For some continuous variables there was evidence ($p<0.05$) that a quadratic term improved the model's fit. In all these cases, however, the underlying relationship was monotonic (i.e. no thresholds or inversions) and no case did the inclusion of this term affect the substantive findings. As such, for simplicity of comparison, I present the results using only the linear terms.

Three risk factors showed a weak/marginal association with externalising score at follow-up which were in the opposite direction to that seen cross-sectionally at baseline. These were smoking ($p=0.07$ for protective effect on parent externalising score); punishment by shouting ($p=0.02$ for a protective effect on parent externalising score); and reward through praising ($p=0.03$ for being a risk factor for the teacher externalising score). I considered these most likely to be chance findings, and therefore classified them as showing no evidence of a predictive effect in the summary grids I present in Chapter 11 (Table 11.6 and Table 11.7).

Parent externalising score

Table 14.46: Full results from unadjusted and adjusted models of putative risk factors (RF) and parent externalising SDQ score (EXT) for binary and ordered categorical variables.

Risk factor	Model	A: RF_{t1} on parent EXT_{t2} (forward causal association): Regression coefficient and 95%CI		B: Parent EXT_{t1} on RF_{t2} (reverse causal association): Odds ratio and 95%CI	
Parental separation N=7261	**Unadjusted**	No Yes	0*** 0.436 (0.283, 0.589)	Change per point	1.045 (1.025, 1.066)***
	Adjusted	No Yes	0** 0.260 (0.096, 0.425)	Change per point	1.042 (1.020, 1.064)***
Financial crisis N=7263	**Unadjusted**	No Yes	0* 0.230 (0.032, 0.429)	Change per point	1.069 (1.041, 1.098)***
	Adjusted	No Yes	0 [p=0.05] 0.195 (-0.001, 0.392)	Change per point	1.046 (1.017, 1.076)**
Family police contact N=7262	**Unadjusted**	No Yes	0** 0.450 (0.130, 0.771)	Change per point	1.066 (1.034, 1.099)***
	Adjusted	No Yes	0 0.212 (-0.112, 0.537)	Change per point	1.054 (1.020, 1.089)**
Death of parent or sibling N=7264	**Unadjusted**	No Yes	0 0.396 (-0.071, 0.862)	Change per point	1.044 (1.021, 1.068)
	Adjusted	No Yes	0 0.272 (-0.204, 0.747)	Change per point	1.027 (1.004, 1.050)*

Risk factor	Model	A: RF_{t1} on parent EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Parent EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Odds ratio and 95%CI	
Neuro-developmental disorder N=7284 (A)/ 2363 (B)	**Unadjusted**	No Yes	0 0.251 (-0.581, 1.082)	Change per point	1.051 (0.854, 1.293)
	Adjusted	No Yes	0 0.325 (-0.488, 1.138)	Change per point	1.006 (0.783, 1.293)
Developmental problems N=7284 (A)/ 2363 (B)	**Unadjusted**	No Yes	0** 0.456 (0.195, 0.716)	Change per point	1.126 (1.072, 1.184)***
	Adjusted	No Yes	0** 0.453 (0.195, 0.710)	Change per point	1.117 (1.059, 1.179)***
Common physical disorder N=7284 (A)/ 2363 (B)	**Unadjusted**	No Yes	0*** 0.223 (0.101, 0.345)	Change per point	1.005 (0.979, 1.032)
	Adjusted	No Yes	0** 0.211 (0.090, 0.332)	Change per point	1.002 (0.974, 1.030
Rare physical disorder N=7284 (A)/ 2363 (B)	**Unadjusted**	No Yes	0* 0.262 (0.013, 0.511)	Change per point	1.040 (0.994, 1.087)
	Adjusted	No Yes	0* 0.270 (0.023, 0.517)	Change per point	1.036 (0.985, 1.089)
Child hospitalisation N=7263	**Unadjusted**	No Yes	0 0.104 (-0.071, 0.278)	Change per point	1.053 (1.025, 1.082)***
	Adjusted	No Yes	0 0.083 (-0.090, 0.256)	Change per point	1.035 (1.004, 1.067)*
Death of friend N=7263	**Unadjusted**	No Yes	0 0.052 (-0.202, 0.307)	Change per point	1.032 (1.006, 1.058)*
	Adjusted	No Yes	0 0.006 (-0.247, 0.259)	Change per point	1.023 (0.997, 1.050) [p=0.09]
Regular smoker N=6284	**Unadjusted**	No Yes	0 -0.420 (-0.987, 0.147)	Change per point	1.179 (1.143, 1.217)***
	Adjusted	No Yes	0 [p=0.07] -0.523 (-1.086, 0.040)	Change per point	1.156 (1.119, 1.195)***
Alcohol consumption N=6281	**Unadjusted**	Never/rare Moderate Frequent	0 -0.259 (-0.543, 0.026) -0.175 (-0.726, 0.377)	Change per point	1.030 (1.008, 1.052)**
	Adjusted	Never/rare Moderate Frequent	0 -0.222 (-0.505, 0.061) -0.165 (-0.733, 0.402)	Change per point	1.046 (1.022, 1.070)**
Ever used drugs N=6279	**Unadjusted**	No Yes	0 -0.123 (-0.589, 0.342)	Change per point	1.082 (1.055, 1.110)***
	Adjusted	No Yes	0 -0.130 (-0.602, 0.342)	Change per point	1.090 (1.063, 1.118)***

Risk factor	Model	A: RF_{t1} on parent EXT_{t2} (forward causal association): Regression coefficient and 95%CI		B: Parent EXT_{t1} on RF_{t2} (reverse causal association): Odds ratio and 95%CI	
Learning difficulty N=7284 (A)/ 2363 (B)	**Unadjusted**	No	0***	Change per point	
		Yes	0.782 (0.465, 1.099)		1.180 (1.124, 1.239)***
	Adjusted	No	0***	Change per point	
		Yes	0.748 (0.433, 1.063)		1.189 (1.128, 1.253)***
Dyslexia N=7284 (A)/2306 (B)	**Unadjusted**	No	0	Change per point	
		Yes	0.424 (0.053, 0.796)*		1.089 (1.021, 1.161)*
	Adjusted	No	0**	Change per point	
		Yes	0.503 (0.131, 0.874)		1.119 (1.043, 1.200)***
Rewards: praise† N= 1750	**Unadjusted**	Never	-0.010 (-2.992, 2.972)	Change per point	
		Seldom	1.518 (0.096, 2.941)		
		Sometimes	0.148 (-0.236, 0.532)		
		Frequently	0		0.925 (0.894, 0.958)***
	Adjusted	Never	0.029 (-2.999, 3.057)	Change per point	
		Seldom	1.173 (-0.046, 2.391)		
		Sometimes	0.050 (-0.326, 0.427)		
		Frequently	0		0.933 (0.899, 0.968)***
Rewards: treats N=1746	**Unadjusted**	Never	0	Change per point	
		Seldom	-0.208 (-0.918, 0.502)		
		Sometimes	0.012 (-0.544, 0.569)		
		Frequently	-0.051 (-0.632, 0.530)		1.043 (1.020, 1.067)***
	Adjusted	Never	0	Change per point	
		Seldom	-0.019 (-0.672, 0.635)		
		Sometimes	0.099 (-0.426, 0.625)		
		Frequently	-0.046 (-0.593, 0.501)		1.033 (1.008, 1.059)**
Rewards: favourite things N= 1734	**Unadjusted**	Never	0	Change per point	
		Seldom	-0.303 (-0.988, 0.382)		
		Sometimes	-0.399 (-0.978, 0.181)		
		Frequently	-0.288 (-0.876, 0.299)		1.033 (1.009, 1.058)**
	Adjusted	Never	0	Change per point	
		Seldom	-0.278 (-0.946, 0.389)		
		Sometimes	-0.396 (-0.970, 0.178)		
		Frequently	-0.343 (-0.927, 0.242)		1.028 (1.002, 1.054)*
Punish: send to room N=1748 [p<0.001 that proportional odds assumption met]	**Unadjusted**	Never	0**	Change per point :	
		Seldom	-0.088 (-0.418, 0.242)	Never vs. above	1.000 (0.967, 1.035)
		Sometimes	-0.191 (-0.567, 0.186)	Seldom vs. above	1.085 (1.055, 1.116)***
		Frequently	0.814 (0.259, 1.370)	Never/seldom/ sometimes vs. above	1.061 (1.014, 1.110)*
	Adjusted	Never	0**	Change per point :	
		Seldom	-0.078 (-0.415, 0.260)	Never vs. above	0.996 (0.961, 1.033)

Risk factor	Model	A: RF_{t1} on parent EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Parent EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Odds ratio and 95%CI	
		Sometimes Frequently	-0.194 (-0.566, 0.178) 0.798 (0.271, 1.324)	Seldom vs. above Never/seldom/ sometimes vs. above	1.075 (1.043, 1.109)*** 1.035 (0.985, 1.087)
Punish: grounding **N=1748** **[p=0.005 that proportional odds assumption met]**	**Unadjusted**	Never Seldom Sometimes Frequently	0** 0.393 (0.098, 0.689) 0.388 (0.057, 0.719) 1.339 (0.514, 2.165)	Change per point : Never vs. above Seldom vs. above Never/seldom/ sometimes vs. above	 1.077 (1.040, 1.116)*** 1.104 (1.068, 1.140)*** 1.119 (1.061, 1.180)***
	Adjusted	Never Seldom Sometimes Frequently	0 [p=0.07] 0.265 (-0.018, 0.548) 0.196 (-0.125, 0.518) 0.946 (0.131, 1.761)	Change per point : Never vs. above Seldom vs. above Never/seldom/ sometimes vs. above	 1.071 (1.031, 1.111)*** 1.094 (1.056, 1.133)*** 1.115 (1.054, 1.179)***
Punish: shouting **N=1750**	**Unadjusted**	Never Seldom Sometimes Frequently	0** -0.501 (-1.138, 0.136) -0.718 (-1.314, -0.122) -0.247 (-0.869, 0.376)	Change per point	1.103 (1.076, 1.131)***
	Adjusted	Never Seldom Sometimes Frequently	0* -0.473 (-1.113, 0.167) -0.652 (-1.248, -0.055) -0.206 (-0.833, 0.421)	Change per point	1.111 (1.082, 1.141)***
Punish: smacking **N=1751** **[p=0.003 that proportional odds assumption met]**	**Unadjusted**	Never Seldom Sometimes Frequently	0*** 0.171 (-0.086, 0.428) -0.127 (-0.593, 0.340) 3.026 (1.820, 4.232)	Change per point Never vs. above Seldom vs. above Never/seldom/ sometimes vs. above	 1.084 (1.050, 1.119)*** 1.168 (1.107, 1.233)*** 1.316 (1.174, 1.475)***
	Adjusted	Never Seldom Sometimes Frequently	0** 0.215 (-0.052, 0.481) -0.079 (-0.552, 0.393) 2.882 (1.314, 4.449)	Change per point Never vs. above Seldom vs. above Never/seldom/ sometimes vs. above	 1.090 (1.054, 1.128)*** 1.163 (1.090, 1.240)*** [Did not converge]
Punish: ever hit or shake **N=1751**	**Unadjusted**	No Yes	0 0.605 (-0.191, 1.401)	Change per point	1.059 (0.979, 1.146)
	Adjusted	No Yes	0 [p=0.06] 0.725 (-0.023, 1.472)	Change per point	1.043 (0.956, 1.137)
Parent disapproval of friends **N=4783**	**Unadjusted**	Approves a lot Approves a little Disapproves	0* 0.394 (0.107, 0.681) -0.295 (-1.259, 0.670)	Change per point	1.147 (1.122, 1.172)***

Risk factor	Model	A: RF_{t1} on parent EXT_{t2} (forward causal association): Regression coefficient and 95%CI		B: Parent EXT_{t1} on RF_{t2} (reverse causal association): Odds ratio and 95%CI	
	Adjusted	Approves a lot Approves a little Disapproves	0* 0.285 (0.008, 0.562) -0.495 (-1.451, 0.460)	Change per point	1.126 (1.101, 1.153)***
Parent thinks friends are trouble **N= 4772** **[p<0.001 that proportional odds assumption met]**	**Unadjusted**	None A few Many All	0*** 0.350 (0.176, 0.524) 0.604 (-0.344, 1.552) 0.993 (-0.981, 2.967)	Change per point: None vs. above None/a few vs. above None/a few/many vs. above	 1.108 (1.089, 1.127)*** 1.216 (1.157, 1.278)*** 1.173 (1.061, 1.297)**
	Adjusted	None A few Many All	0* 0.262 (0.086, 0.437) 0.423 (-0.519, 1.365) 0.944 (-1.008, 2.895)	Change per point: None vs. above None/a few vs. above None/a few/many vs. above	 1.095 (1.075, 1.115)*** 1.192 (1.130, 1.257)*** 1.148 (1.024, 1.287)*

*p<0.05, **p<0.01, ***p<0.001. SD=standard deviation. ‘Unadjusted models’ adjust only for age, gender, survey year and the baseline scores for externalising problems and the risk factor variable in question. ‘Adjusted models’ adjust for these variables and, in addition, ethnic group, parent education, housing tenure, geographical region, metropolitan region, area deprivation, family type, three generation family, number of co-resident siblings, mother’s age at child’s birth. For ordered categorical RF variables as outcomes, proportional odds are presented unless there was evidence (p<0.01) that this model provided a worse fit than non-proportional odds.

† For ‘rewards: praise’, only 8 people were in the bottom category (‘Never’) at baseline, and I therefore present the results with the top category (‘Frequently’) as the baseline

Table 14.47: Full results from unadjusted and adjusted models of putative risk factors (RF) and parent externalising SDQ score (EXT) for continuous variables.

Risk factor	Model	A: RF $_{t1}$ on parent EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Parent EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Regression coefficient and 95%CI		P-value for difference
Parent mental health N=7235	**Unadjusted**	Change per point	0.062 (0.038, 0.086)***	Change per point	0.102 (0.083, 0.120)***	
	Adjusted	Change per point	0.053 (0.029, 0.077)***	Change per point	0.091 (0.071, 0.110)***	
	Adjusted-standardised	Change per SD	0.038 (0.020, 0.055)***	Change per SD	0.118 (0.093, 0.144)***	<0.001
Family functioning N= 7243 (A)/ 2343 (B)	**Unadjusted**	Change per point	0.456 (0.301, 0.611)***	Change per point	0.008 (0.004, 0.012)***	
	Adjusted	Change per point	0.378 (0.222, 0.535)***	Change per point	0.010 (0.005, 0.014)***	
	Adjusted-standardised	Change per SD	0.041 (0.024, 0.058)***	Change per SD	0.085 (0.048, 0.122)***	0.03
Good general health N= 7280	**Unadjusted**	Change per point	-0.243 (-0.351, -0.135)***	Change per point	-0.014 (-0.018, -0.011)***	
	Adjusted	Change per point	-0.194 (-0.303, -0.086)***	Change per point	-0.012 (-0.016, -0.008)***	
	Adjusted-standardised	Change per SD	-0.032 (-0.050, -0.014)***	Change per SD	-0.073 (-0.095, -0.051)***	0.004
Teacher-reported difficulties in school N=2596	**Unadjusted**	Change per point	0.133 (0.083, 0.183)***	Change per point	0.083 (0.066, 0.101)***	
	Adjusted	Change per point	0.103 (0.052, 0.154)***	Change per point	0.071 (0.053, 0.088)***	
	Adjusted-standardised	Change per SD	0.066 (0.033, 0.099)***	Change per SD	0.116 (0.087, 0.144)***	0.03
Parent-reported internalising problems N= 7281	**Unadjusted**	Change per point	0.071 (0.048, 0.094)***	Change per point	0.117 (0.097, 0.136)***	
	Adjusted	Change per point	0.057 (0.035, 0.079)***	Change per point	0.108 (0.088, 0.128)***	
	Adjusted-standardised	Change per SD	0.046 (0.028, 0.064)***	Change per SD	0.129 (0.105, 0.153)***	<0.001
Teacher-reported internalising problems N= 3768	**Unadjusted**	Change per point	0.077 (0.048, 0.106)***	Change per point	0.124 (0.097, 0.152)***	
	Adjusted	Change per point	0.069 (0.040, 0.099)***	Change per point	0.111 (0.084, 0.139)***	
	Adjusted-standardised	Change per SD	0.058 (0.033, 0.082)***	Change per SD	0.136 (0.102, 0.170)***	<0.001
Social aptitudes score N= 4910	**Unadjusted**	Change per point	-0.041 (-0.055, -0.028)***	Change per point	-0.322 (-0.365, -0.279)***	
	Adjusted	Change per point	-0.038 (-0.051, -0.025)***	Change per point	-0.310 (-0.353, -0.267)***	
	Adjusted-standardised	Change per SD	-0.064 (-0.085, -0.042)***	Change per SD	-0.178 (-0.202, -0.153)***	<0.001

*p<0.05, **p<0.01, ***p<0.001. SD=standard deviation. ‘Unadjusted models’ adjust only for age, gender, survey and the baseline scores for externalising problems and the risk factor variable in question.. ‘Adjusted models’ adjust for these variables and, in addition, ethnic group, parent education, housing tenure, geographical region, metropolitan region, area deprivation, family type, three generation family, number of co-resident siblings, mother’s age at child’s birth. ‘Adjusted – standardised’ models standardise both the RF and the EXT variables (i.e. STDXY). P-value for difference is for the difference between the standardised regression coefficients of models A and B, based on their point estimate and standard error.

Teacher externalising score

Table 14.48: Full results from unadjusted and adjusted models of putative risk factors (RF) and teacher externalising SDQ score (EXT) for binary and ordered categorical variables.

Risk factor	Model	A: RF $_{t1}$ on teacher EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Teacher EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Odds ratio and 95%CI	
Parental separation N=3726	**Unadjusted**	No Yes	0*** 0.880 (0.608, 1.151)	Change per point	1.023 (0.993, 1.054)
	Adjusted	No Yes	0*** 0.602 (0.312, 0.892)	Change per point	1.018 (0.988, 1.049)
Financial crisis N=3727	**Unadjusted**	No Yes	0 0.136 (-0.172, 0.444)	Change per point	1.068 (1.026, 1.113)**
	Adjusted	No Yes	0 -0.019 (-0.329, 0.291)	Change per point	1.048 (1.003, 1.095)*
Family police contact N=3727	**Unadjusted**	No Yes	0*** 1.165 (0.522, 1.807)	Change per point	1.042 (0.996, 1.091)
	Adjusted	No Yes	0* 0.737 (0.075, 1.399)	Change per point	1.030 (0.980, 1.082)
Death of parent or sibling N=3727	**Unadjusted**	No Yes	0 -0.349 (-0.986, 0.288)	Change per point	1.047 (1.004, 1.093)*
	Adjusted	No Yes	0* -0.677 (-1.300, -0.053)	Change per point	1.034 (0.989, 1.080)
Neuro-developmental disorder N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0 1.678 (-0.359, 3.714)	Change per point	1.017 (0.774, 1.335)
	Adjusted	No Yes	0 1.630 (-0.350, 3.609)	Change per point	0.960 (0.699, 1.319)
Developmental problems N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0*** 0.728 (0.371, 1.084)	Change per point	1.125 (1.062, 1.191)***
	Adjusted	No Yes	0*** 0.655 (0.306, 1.004)	Change per point	1.120 (1.047, 1.198)**
Common physical disorder N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0* 0.222 (0.005, 0.438)	Change per point	1.000 (0.968, 1.034)
	Adjusted	No Yes	0 [p=0.09] 0.183 (-0.031, 0.397)	Change per point	0.996 (0.962, 1.032)
Rare physical disorder N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0 0.318 (-0.137, 0.773)	Change per point	1.016 (0.963, 1.072)
	Adjusted	No Yes	 0.265 (-0.187, 0.716)	Change per point	1.007 (0.948, 1.070)

Risk factor	Model	A: RF $_{t1}$ on teacher EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Teacher EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Odds ratio and 95%CI	
Child hospitalisation N= 3726	**Unadjusted**	No Yes	0* 0.342 (0.016, 0.667)	Change per point	 1.076 (1.038, 1.115)***
	Adjusted	No Yes	0 [p=0.07] 0.295 (-0.028, 0.618)	Change per point	 1.063 (1.024, 1.103)**
Death of friend N=3725	**Unadjusted**	No Yes	0 -0.191 (-0.631, 0.250)	Change per point	 1.045 (1.011, 1.080)*
	Adjusted	No Yes	0 -0.179 (-0.603, 0.245)	Change per point	 1.037 (1.000, 1.076)*
Regular smoker N= 3417	**Unadjusted**	No Yes	0 1.452 (-0.637, 3.542)	Change per point	 1.187 (1.131, 1.245)***
	Adjusted	No Yes	0 1.570 (-0.573, 3.712)	Change per point	 1.164 (1.104, 1.228)***
Alcohol consumption N= 3415	**Unadjusted**	Never/rare Moderate Frequent	0 0.279 (-0.462, 1.020) -0.175 (-1.751, 1.401)	Change per point	 1.052 (1.018, 1.086)**
	Adjusted	Never/rare Moderate Frequent	0 0.328 (-0.394, 1.051) -0.169 (-1.660, 1.321)	Change per point	 1.067 (1.028, 1.106)**
Ever used drugs N=3414	**Unadjusted**	No Yes	0 0.655 (-0.921, 2.230)	Change per point	 1.096 (1.055, 1.137)***
	Adjusted	No Yes	0 0.813 (-0.753, 2.379)	Change per point	 1.088 (1.045, 1.133)***
Learning difficulty N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0*** 1.448 (0.897, 2.000)	Change per point	 1.193 (1.126, 1.264)***
	Adjusted	No Yes	0*** 1.335 (0.797, 1.874)	Change per point	 1.196 (1.129, 1.267)***
Dyslexia N= 3739(A)/ 1395 (B)	**Unadjusted**	No Yes	0 0.452 (-0.295, 1.199)	Change per point	 1.136 (1.055, 1.224)**
	Adjusted	No Yes	0 0.512 (-0.203, 1.228)	Change per point	 1.205 (1.102, 1.318)***
Rewards: praise† N= 1182	**Unadjusted**	Never Seldom Sometimes Frequently	-0.862 (-2.188, 0.463) -0.979 (-2.168, 0.210) -0.074 (-0.574, 0.427) 0	Change per point	 0.970 (0.930, 1.012)
	Adjusted	Never Seldom Sometimes Frequently	-1.122 (-2.553, 0.308) -1.391 (-2.439, -0.342) -0.133 (-0.630, 0.365) 0	Change per point	 0.986 (0.941, 1.034)

Risk factor	Model	A: RF_{t1} on teacher EXT_{t2} (forward causal association): Regression coefficient and 95%CI		B: Teacher EXT_{t1} on RF_{t2} (reverse causal association): Odds ratio and 95%CI	
Rewards: treats N=1179	**Unadjusted**	Never	0 [p=0.07]	Change per point	
		Seldom	-0.126 (-1.110, 0.858)		
		Sometimes	-0.230 (-1.057, 0.597)		
		Frequently	0.329 (-0.526, 1.184)		1.058 (1.024, 1.094)**
	Adjusted	Never	0	Change per point	
		Seldom	-0.013 (-1.058, 1.033)		
		Sometimes	-0.141 (-1.047, 0.766)		
		Frequently	0.311 (-0.641, 1.262)		1.051 (1.011, 1.092)*
Rewards: favourite things N= 1170	**Unadjusted**	Never	0	Change per point	
		Seldom	0.049 (-0.883, 0.982)		
		Sometimes	-0.318 (-1.078, 0.442)		
		Frequently	0.142 (-0.637, 0.920)		1.023 (0.994, 1.053)
	Adjusted	Never	0	Change per point	
		Seldom	-0.058 (-1.007, 0.890)		
		Sometimes	-0.353 (-1.160, 0.454)		
		Frequently	-0.008 (-0.847, 0.830)		1.020 (0.986, 1.055)
Punish: send to room N=1181	**Unadjusted**	Never	0*	Change per point :	
		Seldom	-0.121 (-0.587, 0.346)		
		Sometimes	0.216 (-0.231, 0.662)		
		Frequently	1.264 (0.371, 2.157)		1.060 (1.027, 1.093)***
	Adjusted	Never	0*	Change per point :	
		Seldom	-0.054 (-0.520, 0.412)		
		Sometimes	0.289 (-0.161, 0.739)		
		Frequently	1.192 (0.339, 2.045)		1.054 (1.018, 1.091)**
Punish: grounding N=1181 [p=0.03 that proportional odds assumption met]	**Unadjusted**	Never	0***	Change per point :	
		Seldom	0.101 (-0.368, 0.571)	Never vs. above	1.075 (1.034, 1.118)***
		Sometimes	0.557 (0.075, 1.039)	Seldom vs. above	1.120 (1.077, 1.165)***
		Frequently	2.686 (1.289, 4.084)	Never/seldom/ sometimes vs. above	1.184 (1.115, 1.257)***
	Adjusted	Never	0**	Change per point :	
		Seldom	-0.042 (-0.519, 0.435)	Never vs. above	1.070 (1.025, 1.117)**
		Sometimes	0.360 (-0.120, 0.840)	Seldom vs. above	1.116 (1.070, 1.164)***
		Frequently	2.317 (0.952, 3.682)	Never/seldom/ sometimes vs. above	1.193 (1.125, 1.266)***
Punish: shouting N= 1182	**Unadjusted**	Never	0	Change per point	
		Seldom	-0.086 (-1.227, 1.056)		
		Sometimes	-0.102 (-1.149, 0.945)		
		Frequently	0.275 (-0.793, 1.343)		1.040 (1.010, 1.071)**
	Adjusted	Never	0	Change per point	

Risk factor	Model	A: RF $_{t1}$ on teacher EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Teacher EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Odds ratio and 95%CI	
		Seldom	0.092 (-1.085, 1.268)		
		Sometimes	0.092 (-0.967, 1.152)		
		Frequently	0.500 (-0.590, 1.590)		1.048 (1.015, 1.082)**
Punish: smacking N=1183	**Unadjusted**	Never	0 [p=0.07]	Change per point	
		Seldom	0.509 (0.115, 0.903)		
		Sometimes	0.169 (-0.445, 0.784)		
		Frequently	1.057 (-1.212, 3.327)		1.076 (1.036, 1.117)***
	Adjusted	Never	0*	Change per point	
		Seldom	0.649 (0.245, 1.053)		
		Sometimes	0.224 (-0.387, 0.834)		
		Frequently	0.445 (-1.824, 2.714)		1.077 (1.032, 1.124)**
Punish: ever hit or shake N=1183	**Unadjusted**	No		Change per point	
		Yes	0.807 (-0.205, 1.818)		1.043 (0.914, 1.189)
	Adjusted	No		Change per point	
		Yes	0.995 (0.024, 1.965)*		1.046 (0.917, 1.194)
Parent disapproval of friends N=2297	**Unadjusted**	Approves a lot	0*	Change per point	
		Approves a little	0.649 (0.204, 1.094)		
		Disapproves	0.137 (-1.399, 1.674)		1.122 (1.090, 1.155)***
	Adjusted	Approves a lot	0 [p=0.09]	Change per point	
		Approves a little	0.496 (0.043, 0.949)		
		Disapproves	-0.223 (-1.872, 1.427)		1.104 (1.072, 1.137)***
Parent thinks friends are trouble N=2290	**Unadjusted**	None	0***	Change per point:	
		A few	0.462 (0.174, 0.750)		
		Many	0.987 (-1.203, 3.178)		1.073 (1.044, 1.102)***
		All	-0.600 (-1.201, 0.000)		
	Adjusted	None	0	Change per point:	
		A few	0.325 (0.040, 0.611)		
		Many	0.744 (-1.554, 3.042)		
		All	-0.080 (-0.853, 0.694)		1.055 (1.025, 1.085)***

*p<0.05, **p<0.01, ***p<0.001. SD=standard deviation. 'Unadjusted models' adjust only for age, gender, survey year and the baseline scores for externalising problems and the risk factor variable in question. 'Adjusted models' adjust for these variables and, in addition, ethnic group, parent education, housing tenure, geographical region, metropolitan region, area deprivation, family type, three generation family, number of co-resident siblings, mother's age at child's birth. For ordered categorical RF variables as outcomes, proportional odds are presented unless there was evidence (p<0.01) that this model provided a worse fit than non-proportional odds.

† For 'rewards: praise', only 4 people were in the bottom category ('Never') at baseline, and I therefore present the results with the top category ('Frequently') as the baseline.

Table 14.49: Full results from models of putative risk factors (RF) and teacher externalising SDQ score (EXT) for continuous variables.

Risk factor	Model	A: RF $_{t1}$ on teacher EXT $_{t2}$ (forward causal association): Regression coefficient and 95%CI		B: Teacher EXT $_{t1}$ on RF $_{t2}$ (reverse causal association): Regression coefficient and 95%CI		P-value for difference
Parent mental health N=3176	**Unadjusted**	Change per point	0.053 (0.013, 0.092)**	Change per point	0.039 (0.015, 0.063)**	
	Adjusted	Change per point	0.031 (-0.009, 0.070)	Change per point	0.028 (0.004, 0.053)*	
	Adjusted-standardised	Change per SD	0.021 (-0.006, 0.048)	Change per SD	0.038 (0.005, 0.072)*	0.44
Family functioning N= 3728 (A)/ 1389 (B)	**Unadjusted**	Change per point	0.354 (0.095, 0.612)**	Change per point	0.003 (-0.003, 0.008)	
	Adjusted	Change per point	0.255 (-0.015, 0.526)*	Change per point	0.002 (-0.003, 0.007)	
	Adjusted-standardised	Change per SD	0.027 (-0.002, 0.055)*	Change per SD	0.019 (-0.029, 0.068)	0.80
Good general health N= 3735	**Unadjusted**	Change per point	-0.340 (-0.516, -0.164)***	Change per point	-0.006 (-0.012, -0.001)*	
	Adjusted	Change per point	-0.263 (-0.442, -0.084)**	Change per point	-0.004 (-0.009, 0.001)	
	Adjusted-standardised	Change per SD	-0.042 (-0.071, -0.014)**	Change per SD	-0.026 (-0.059, 0.008)	0.46
Teacher-reported difficulties in school N=2537	**Unadjusted**	Change per point	0.205 (0.147, 0.264)***	Change per point	0.054 (0.032, 0.076)***	
	Adjusted	Change per point	0.168 (0.110, 0.226)***	Change per point	0.045 (0.024, 0.066)***	
	Adjusted-standardised	Change per SD	0.105 (0.069, 0.142)***	Change per SD	0.078 (0.041, 0.114)***	0.28
Parent-reported internalising problems N= 3739	**Unadjusted**	Change per point	0.063 (0.026, 0.100)**	Change per point	0.089 (0.063, 0.115)***	
	Adjusted	Change per point	0.037 (-0.001, 0.075) [p=0.05]	Change per point	0.079 (0.052, 0.106)***	
	Adjusted-standardised	Change per SD	0.029 (0.000, 0.059) [p=0.05]	Change per SD	0.098 (0.065, 0.132)***	0.002
Teacher-reported internalising problems N= 3735	**Unadjusted**	Change per point	-0.023 (-0.059, 0.013)	Change per point	0.102 (0.071, 0.132)***	
	Adjusted	Change per point	-0.032 (-0.068, 0.004) [p=0.08]	Change per point	0.085 (0.055, 0.116)***	
	Adjusted-standardised	Change per SD	-0.026 (-0.055, 0.003) [p=0.08]	Change per SD	0.109 (0.070, 0.148)***	<0.001
Social aptitudes score N= 2334	**Unadjusted**	Change per point	-0.038 (-0.060, -0.016)**	Change per point	-0.254 (-0.318, -0.190)***	
	Adjusted	Change per point	-0.034 (-0.057, -0.012)**	Change per point	-0.231 (-0.293, -0.168)***	
	Adjusted-standardised	Change per SD	-0.055 (-0.092, -0.019)**	Change per SD	-0.138 (-0.175, -0.101)***	0.002

*p<0.05, **p<0.01, ***p<0.001. SD=standard deviation. 'Unadjusted models' adjust only for age, gender, survey year and the baseline scores for externalising problems and the risk factor variable in question. 'Adjusted models' adjust for these variables and, in addition, ethnic group, parent education, housing tenure, geographical region, metropolitan region, area deprivation, family type, three generation family, number of co-resident siblings, mother's age at child's birth. 'Adjusted – standardised' models standardise both the RF and the EXT variables (i.e. STDXY). P-value for difference is for the difference between the standardised regression coefficients of models A and B, based on their point estimate and standard error.

14.8.3 Full models of selected final multivariable analyses

Table 14.50: Fully adjusted models predicting to parent-reported externalising problems in final multivariable analyses (N=14 229; 13 868 White, 361 Indian)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
Ethnicity	**Ethnicity**	Indian	0***	0***	0***	0***	0***
		White	0.98 (0.61, 1.34)	1.07 (0.71, 1.43)	0.71 (0.35, 1.06)	1.04 (0.67, 1.42)	0.75 (0.38, 1.11)
A priori confounders	**Child's sex**	Male	0***	0***	0***	0***	0***
		Female	-1.35 (-1.47, -1.24)	-1.33 (-1.44, -1.21)	-0.89 (-1.01, -0.78)	-1.31 (-1.42, -1.20)	-0.93 (-1.04, -0.82)
	Child's age	Change per year	-0.10 (-0.12, -0.08)***	-0.15 (-0.16, -0.13)***	-0.05 (-0.07, -0.03)***	-0.14 (-0.16, -0.12)***	-0.10 (-0.12, -0.08)***
	Survey year	1999	0	0*	0*	0	0
		2004	-0.08 (-0.24, 0.07)	-0.18 (-0.32, -0.04)	-0.16 (-0.30, -0.02)	-0.06 (-0.21, 0.09)	-0.07 (-0.21, 0.07)
Area	**Geographical region**	South East	0			0	0
		London	-0.16 (-0.43, 0.11)			-0.14 (-0.40, 0.12)	0.00 (-0.25, 0.26)
		South West	-0.16 (-0.43, 0.11)			-0.22 (-0.47, 0.03)	-0.09 (-0.33, 0.15)
		Eastern	-0.23 (-0.48, 0.02)			-0.24 (-0.47, 0.00)	-0.14 (-0.36, 0.08)
		East Midlands	-0.09 (-0.40, 0.21)			-0.12 (-0.42, 0.18)	0.06 (-0.22, 0.34)
		West Midlands	-0.07 (-0.37, 0.24)			-0.05 (-0.34, 0.24)	0.08 (-0.18, 0.33)
		North East	-0.11 (-0.52, 0.30)			0.03 (-0.36, 0.42)	0.20 (-0.18, 0.57)
		North West & Merseyside	-0.16 (-0.40, 0.08)			-0.10 (-0.34, 0.14)	0.05 (-0.17, 0.28)
		Yorkshire & Humberside	-0.09 (-0.35, 0.18)			-0.09 (-0.35, 0.18)	0.10 (-0.15, 0.35)
	Metropolitan region	Non-Metropolitan	0			0	0
		Metropolitan	-0.05 (-0.21, 0.12)			-0.02 (-0.18, 0.14)	-0.01 (-0.16, 0.14)
	Area deprivation	Change per standard deviation	0.03 (-0.06, 0.13)			-0.02 (-0.12, 0.07)	0.01 (-0.08, 0.10)
School	**Ford score**	Change per point	0.09 (0.07, 0.12)***			0.08 (0.05, 0.10)***	0.05 (0.03, 0.08)***
Family SEP	**Parent's highest educational qualification**	No qualifications	0***			0***	0***
		Poor GCSEs	-0.21 (-0.42, 0.01)			-0.21 (-0.42, 0.00)	-0.08 (-0.28, 0.13)
		Good GCSEs	-0.70 (-0.91, -0.50)			-0.61 (-0.81, -0.42)	-0.27 (-0.45, -0.09)
		A-level	-0.89 (-1.14, -0.64)			-0.76 (-1.01, -0.52)	-0.33 (-0.55, -0.10)
		Diploma	-0.97 (-1.22, -0.72)			-0.81 (-1.04, -0.57)	-0.42 (-0.64, -0.20)
		Degree	-1.40 (-1.66, -1.13)			-1.19 (-1.44, -0.93)	-0.61 (-0.84, -0.38)
	Weekly household Income	£0-99	0**			0	0
		£100-199	0.16 (-0.27, 0.59)			0.08 (-0.34, 0.51)	0.10 (-0.28, 0.48)
		£200-299	0.12 (-0.31, 0.56)			0.13 (-0.30, 0.56)	0.15 (-0.24, 0.55)
		£300-399	0.08 (-0.35, 0.52)			0.18 (-0.26, 0.62)	0.25 (-0.16, 0.67)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
		£400-499	0.02 (-0.43, 0.47)			0.20 (-0.27, 0.66)	0.30 (-0.12, 0.73)
		£500-599	-0.22 (-0.66, 0.22)			-0.01 (-0.48, 0.46)	0.20 (-0.23, 0.64)
		£600-769	-0.30 (-0.73, 0.13)			-0.02 (-0.47, 0.43)	0.22 (-0.20, 0.63)
		£770 and over	-0.32 (-0.74, 0.10)			0.04 (-0.40, 0.48)	0.32 (-0.09, 0.73)
	Housing tenure	Owner occupied	0***			0***	0***
		Social sector rented	1.10 (0.89, 1.31)			0.70 (0.50, 0.91)	0.37 (0.18, 0.57)
		Privately rented	0.67 (0.38, 0.95)			0.40 (0.13, 0.67)	0.28 (0.03, 0.54)
	Occupational social class	I	0**			0 [p=0.09]	0
		II	0.04 (-0.23, 0.32)			-0.01 (-0.27, 0.25)	-0.12 (-0.36, 0.13)
		III Non-manual	0.30 (0.01, 0.60)			0.17 (-0.11, 0.46)	0.11 (-0.16, 0.37)
		III Manual	0.23 (-0.07, 0.53)			0.12 (-0.17, 0.41)	-0.08 (-0.36, 0.20)
		IV	0.39 (0.08, 0.70)			0.22 (-0.08, 0.52)	0.03 (-0.25, 0.32)
		V	0.52 (0.11, 0.93)			0.29 (-0.11, 0.68)	0.01 (-0.37, 0.39)
		Never worked	1.14 (0.54, 1.73)			0.72 (0.14, 1.31)	0.48 (-0.08, 1.05)
		Full-time student	0.40 (-0.31, 1.11)			0.36 (-0.33, 1.04)	-0.04 (-0.70, 0.61)
Family composition	**Family type**	Two-parent family		0***		0***	0***
		Step family		0.83 (0.58, 1.08)		0.60 (0.35, 0.85)	0.47 (0.23, 0.70)
		Lone parent family		0.59 (0.34, 0.84)		0.06 (-0.21, 0.34)	0.06 (-0.20, 0.32)
	Three generation family	No grandparent in household		0		0	0
		Grandparent in household		0.17 (-0.24, 0.58)		-0.03 (-0.44, 0.38)	-0.16 (-0.55, 0.22)
	Number of co-resident siblings	0		0***		0 [p=0.06]	0
		1		-0.01 (-0.17, 0.15)		0.13 (-0.03, 0.30)	0.09 (-0.07, 0.24)
		2		-0.01 (-0.20, 0.17)		0.07 (-0.12, 0.25)	0.02 (-0.15, 0.20)
		3		0.57 (0.29, 0.85)		0.39 (0.11, 0.67)	0.24 (-0.02, 0.50)
		4 or more		0.69 (0.24, 1.15)		0.34 (-0.12, 0.79)	0.00 (-0.41, 0.42)
	Mother's age at child's birth	Change per decade		-0.95 (-1.08, -0.81)		-0.05 (-0.07, -0.04)***	-0.05 (-0.06, -0.04)***
Family stress	**Family functioning**	Change per standard deviation		0.92 (0.86, 0.99)***		0.82 (0.76, 0.89)***	0.75 (0.68, 0.81)***
	Parental separation	No		0**		0**	0**
		Yes		0.37 (0.13, 0.60)		0.40 (0.17, 0.63)	0.34 (0.11, 0.56)
	Family financial crisis	No		0*		0*	0
		Yes		0.23 (0.03, 0.42)		0.25 (0.05, 0.44)	0.15 (-0.03, 0.32)
	Family police contact	No		0***		0***	0***
		Yes		0.88 (0.58, 1.17)		0.65 (0.36, 0.94)	0.52 (0.24, 0.79)
	Death of parent or sibling	No		0*		0*	0
		Yes		0.40 (0.04, 0.75)		0.37 (0.01, 0.73)	0.17 (-0.16, 0.50)
Child	**Neuro-developmental disorder**	No			0 [p=0.06]		0 [p=0.05]
		Yes			0.66 (-0.03, 1.35)		0.62 (0.00, 1.24)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
	Developmental problems	No Yes			0*** 1.09 (0.85, 1.33)		0*** 0.94 (0.71, 1.17)
	Common physical disorder	No Yes			0*** 0.45 (0.33, 0.57)		0*** 0.39 (0.27, 0.50)
	Rare physical disorder	No Yes			0** 0.33 (0.08, 0.58)		0** 0.35 (0.11, 0.58)
	Child hospitalisation	No Yes			0* 0.21 (0.04, 0.38)		0* 0.16 (0.00, 0.32)
	Death of friend	No Yes			0*** 0.62 (0.34, 0.90)		0*** 0.55 (0.28, 0.81)
	Teacher-reported academic difficulties	Change per point			0.48 (0.45, 0.50)***		0.34 (0.32, 0.37)***
	Learning difficulty	No Yes			0*** 1.68 (1.41, 1.96)		0* 1.67 (1.42, 1.92)
	Dyslexia	No Yes			0 -0.06 (-0.42, 0.30)		0* 0.45 (0.11, 0.78)

Table 14.51: Fully adjusted models predicting to teacher-reported externalising problems in final multivariable analyses (N=11 032; 10 775 White, 257 Indian)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
Ethnicity	**Ethnicity**	Indian White	0*** 0.93 (0.53, 1.33)	0*** 0.91 (0.51, 1.31)	0*** 0.76 (0.42, 1.10)***	0*** 0.88 (0.45, 1.30)	0*** 0.70 (0.31, 1.08)
A priori confounders	**Child's sex**	Male Female	0*** -2.08 (-2.21, -1.94)	0*** -2.08 (-2.22, -1.94)	0*** -1.64 (-1.77, -1.51)	0*** -2.07 (-2.20, -1.93)	0*** -1.67 (-1.79, -1.54)
	Child's age	Change per year	-0.05 (-0.08, -0.03)***	-0.09 (-0.12, -0.07)***	0.01 (-0.02, 0.03)	-0.09 (-0.11, -0.06)***	-0.03 (-0.05, -0.01)*
	Survey year	1999 2004	0 0.08 (-0.07, 0.24)	0 0.00 (-0.13, 0.14)	0 0.06 (-0.07, 0.19)	0 0.09 (-0.06, 0.24)	0 0.06 (-0.08, 0.20)
Area	**Geographical region**	South East London South West Eastern East Midlands	0 -0.24 (-0.55, 0.07) -0.06 (-0.31, 0.19) -0.18 (-0.44, 0.08) -0.33 (-0.61, -0.05)			0 -0.23 (-0.55, 0.08) -0.10 (-0.34, 0.14) -0.21 (-0.47, 0.05) -0.36 (-0.63, -0.08)	0 -0.04 (-0.33, 0.25) 0.02 (-0.22, 0.25) -0.09 (-0.32, 0.14) -0.15 (-0.40, 0.11)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
		West Midlands	-0.12 (-0.39, 0.16)			-0.13 (-0.41, 0.15)	-0.02 (-0.27, 0.24)
		North East	-0.40 (-0.72, -0.09)			-0.34 (-0.65, -0.03)	-0.21 (-0.52, 0.11)
		North West & Merseyside	-0.22 (-0.47, 0.02)			-0.22 (-0.46, 0.02)	-0.09 (-0.33, 0.14)
		Yorkshire & Humberside	-0.29 (-0.58, -0.01)			-0.30 (-0.58, -0.02)	-0.13 (-0.39, 0.12)
	Metropolitan region	Non-Metropolitan	0			0	0
		Metropolitan	0.10 (-0.08, 0.27)			0.12 (-0.06, 0.29)	0.12 (-0.04, 0.28)
	Area deprivation	Change per standard deviation	0.03 (-0.07, 0.13)			-0.02 (-0.12, 0.08)	0.01 (-0.08, 0.10)
School	**Ford score**	Change per point	0.04 (0.01, 0.07)**			0.04 (0.01, 0.06)*	0.01 (-0.02, 0.03)
Family SEP	**Parent's highest educational qualification**	No qualifications	0***			0***	0
		Poor GCSEs	-0.48 (-0.76, -0.20)			-0.47 (-0.75, -0.19)	-0.25 (-0.52, 0.01)
		Good GCSEs	-0.87 (-1.11, -0.63)			-0.84 (-1.08, -0.61)	-0.35 (-0.59, -0.11)
		A-level	-0.99 (-1.28, -0.70)			-0.95 (-1.24, -0.66)	-0.27 (-0.55, 0.02)
		Diploma	-0.99 (-1.28, -0.69)			-0.91 (-1.20, -0.61)	-0.25 (-0.53, 0.02)
		Degree	-1.21 (-1.49, -0.92)			-1.11 (-1.39, -0.82)	-0.21 (-0.48, 0.07)
	Weekly household income	£0-99	0***			0 [p=0.06]	0**
		£100-199	-0.26 (-0.74, 0.23)			-0.23 (-0.71, 0.26)	-0.21 (-0.66, 0.24)
		£200-299	-0.27 (-0.78, 0.24)			-0.10 (-0.63, 0.43)	-0.02 (-0.51, 0.46)
		£300-399	-0.62 (-1.12, -0.12)			-0.35 (-0.88, 0.18)	-0.19 (-0.68, 0.30)
		£400-499	-0.82 (-1.33, -0.30)			-0.46 (-1.00, 0.08)	-0.29 (-0.78, 0.21)
		£500-599	-0.96 (-1.47, -0.44)			-0.54 (-1.10, 0.02)	-0.26 (-0.79, 0.26)
		£600-769	-0.93 (-1.41, -0.45)			-0.47 (-1.01, 0.06)	-0.12 (-0.62, 0.39)
		£770 and over	-0.76 (-1.25, -0.28)			-0.23 (-0.77, 0.30)	0.16 (-0.34, 0.67)
	Housing tenure	Owner occupied	0***			0***	0**
		Social sector rented	1.16 (0.91, 1.40)			0.85 (0.60, 1.09)	0.42 (0.18, 0.65)
		Privately rented	0.58 (0.25, 0.92)			0.32 (-0.01, 0.65)	0.20 (-0.11, 0.50)
	Occupational social class	I	0			0	0
		II	0.32 (0.04, 0.60)			0.26 (-0.02, 0.54)	0.10 (-0.15, 0.34)
		III Non-manual	0.48 (0.16, 0.80)			0.36 (0.04, 0.68)	0.22 (-0.06, 0.50)
		III Manual	0.42 (0.08, 0.75)			0.37 (0.03, 0.70)	0.03 (-0.26, 0.33)
		IV	0.54 (0.18, 0.89)			0.43 (0.08, 0.78)	0.13 (-0.18, 0.44)
		V	0.63 (0.14, 1.13)			0.48 (-0.01, 0.96)	-0.01 (-0.46, 0.44)
		Never worked	0.89 (0.10, 1.69)			0.65 (-0.17, 1.46)	0.22 (-0.56, 1.00)
		Full-time student	0.58 (-0.27, 1.43)			0.47 (-0.36, 1.31)	-0.02 (-0.78, 0.73)
Family composition	**Family type**	Two-parent family		0***		0**	0*
		Step family		0.63 (0.35, 0.91)		0.44 (0.16, 0.73)	0.33 (0.07, 0.58)
		Lone parent family		0.77 (0.47, 1.07)		0.16 (-0.18, 0.49)	0.16 (-0.16, 0.47)
	Three generation family	No grandparent in household		0*		0	0
		Grandparent in		0.55 (0.00, 1.10)		0.43 (-0.11, 0.96)	0.22 (-0.26, 0.70)

Domain	Variable	Categories	Level 1	Level 2	Level 3	Levels 1 and 2	All levels
		household					
	Number of co-resident siblings	0 1 2 3 4 or more		0*** -0.29 (-0.48, -0.09) -0.37 (-0.60, -0.15) 0.45 (0.11, 0.79) 0.68 (0.13, 1.22)		0*** -0.18 (-0.37, 0.02) -0.33 (-0.55, -0.11) 0.25 (-0.09, 0.58) 0.30 (-0.25, 0.84)	0** -0.22 (-0.40, -0.04) -0.38 (-0.58, -0.18) 0.03 (-0.29, 0.34) -0.10 (-0.61, 0.40)
	Mother's age at child's birth	Change per decade		-0.68 (-0.83, -0.53)***		-0.04 (-0.05, -0.02)***	-0.03 (-0.05, -0.02)***
Family stress	**Family functioning**	Change per standard deviation		0.41 (0.34, 0.49)***		0.31 (0.23, 0.39)***	0.24 (0.17, 0.30)***
	Parental separation	No Yes		0*** 0.58 (0.33, 0.84)		0*** 0.60 (0.34, 0.85)	0*** 0.12 (-0.06, 0.31)
	Family financial crisis	No Yes		0 0.16 (-0.04, 0.36)		0 0.17 (-0.03, 0.37)	0 0.12 (-0.06, 0.31)
	Family police contact	No Yes		0*** 0.73 (0.35, 1.10)		0* 0.47 (0.10, 0.84)	0* 0.42 (0.09, 0.75)
	Death of parent or sibling	No Yes		0** 0.75 (0.30, 1.19)		0** 0.68 (0.23, 1.12)	0* 0.50 (0.25, 0.74)
Child	**Neuro-developmental disorder**	No Yes			0 -0.15 (-0.93, 0.64)		0 -0.19 (-0.92, 0.54)
	Developmental problems	No Yes			0 0.08 (-0.19, 0.35)		0 0.02 (-0.25, 0.28)
	Common physical disorder	No Yes			0* 0.15 (0.02, 0.27)		0 0.10 (-0.02, 0.22)
	Rare physical disorder	No Yes			0 -0.08 (-0.37, 0.20)		0 -0.04 (-0.33, 0.24)
	Child hospitalisation	No Yes			0 0.05 (-0.13, 0.24)		0 0.00 (-0.18, 0.19)
	Death of friend	No Yes			0* 0.36 (0.07, 0.64)		0 [p=0.07] 0.26 (-0.02, 0.54)
	Teacher-reported academic difficulties	Change per point			0.69 (0.66, 0.72)***		0.61 (0.58, 0.65)***
	Learning difficulty	No Yes			0** 0.62 (0.27, 0.96)		0*** 0.62 (0.28, 0.96)
	Dyslexia	No Yes			0*** -0.77 (-1.19, -0.34)		0* -0.50 (-0.92, -0.08)

14.8.4 Replication with alternative mental health outcomes

Table 14.52: Repeating the final multivariable analyses with the DAWBA band

		Proportional odds ratio from logistic regression		
	Adjusted for:	**Parent DAWBA (13 868 White, 361 Indian)**	**Teacher DAWBA (10 775 White, 257 Indian)**	**Child DAWBA (5737 White, 154 Indian)**
Behavioural DAWBA band	**Sex, age and survey year**	1.97 (1.60, 2.42)***	1.61 (1.21, 2.14)**	1.65 (1.19, 2.28)**
	Plus academic difficulties and learning difficulties	1.86 (1.50, 2.30)***	1.53 (1.13, 2.06)**	1.64 (1.19, 2.27)**
	Plus family type and parental divorce	1.70 (1.37, 2.11)***	1.31 (0.97, 1.77) [p=0.08]	1.41 (1.02, 1.94)*
	Plus area, school and family SEP	1.78 (1.43, 2.23)***	1.44 (1.06, 1.97)*	1.45 (1.04, 2.02)*
	Plus other family composition and stress	1.80 (1.43, 2.26)***	1.47 (1.07, 2.03)*	1.48 (1.07, 2.06)*
	Plus other child variables	1.75 (1.39, 2.19)***	1.47 (1.07, 2.03)	1.46 (1.05, 2.03)*
	Plus; family functioning	2.03 (1.63, 2.53)***	1.53 (1.11, 2.11)**	1.58 (1.14, 2.21)**
Hyperactivity DAWBA band	**Sex, age and survey year**	3.23 (2.08, 5.00)***	1.57 (1.09, 2.27)*	–
	Plus academic difficulties and learning difficulties	2.79 (1.71, 4.55)***	1.43 (0.98, 2.09) [p=0.07]	–
	Plus family type and parental divorce	2.25 (1.40, 3.62)**	1.26 (0.86, 1.85)	–
	Plus area, school and family SEP	2.46 (1.49, 4.04)***	1.38 (0.92, 2.07)	–
	Plus other family composition and stress	2.44 (1.48, 4.03)***	1.39 (0.92, 2.08)	–
	Plus other child variables	2.21 (1.33, 3.65)**	1.39 (0.93, 2.09)	–
	Plus; family functioning	2.37 (1.41, 3.98)**	1.43 (0.96, 2.15) [p=0.08]	–
Emotional DAWBA band	**Sex, age and survey year**	1.30 (0.99, 1.70) [p=0.06]	1.59 (0.78, 3.23)	1.52 (0.94, 2.47) [p=0.09]
	Plus academic difficulties and learning difficulties	1.18 (0.90, 1.55)	1.36 (0.64, 2.89)	1.49 (0.92, 2.40)
	Plus family type and parental divorce	1.02 (0.78, 1.33)	1.22 (0.57, 2.59)	1.30 (0.81, 2.10)
	Plus area, school and family SEP	1.09 (0.83, 1.43)	1.37 (0.63, 2.98)	1.48 (0.91, 2.41)
	Plus other family composition and stress	1.09 (0.83, 1.44)	1.41 (0.65, 3.07)	1.50 (0.92, 2.44)
	Plus other child variables	0.98 (0.75, 1.30)	1.37 (0.63, 2.96)	1.43 (0.89, 2.32)
	Plus; family functioning	1.06 (0.80, 1.39)	1.42 (0.66, 3.05)	1.49 (0.92, 2.40)

*p<0.05, **p<0.01, ***p<0.001. Values in the table are regression coefficients for White (vs. Indian) ethnicity from ordered logistic regression. There is no DAWBA band for hyperactivity disorders by child report.

14.8.5 Interactions between ethnicity and family SEP

Table 14.53: Effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity; stratified analyses by parent education

	Parent externalising score					**Teacher externalising score**				
Adjusted for:	**Full population (13 815 White, 358 Indian)**	**p-value for interaction with parent education**	**A-level qualifications or above (4698 White, 124 Indian)**	**GCSE-level qualifications (6400 White, 132 Indian)**	**No education (2717 White, 102 Indian)**	**Full population (10,752 White, 255 Indian)**	**p-value for interaction with parent education**	**A-level qualifications or above (3752 White, 95 Indian)**	**GCSE-level qualifications (4990 White, 98 Indian)**	**No education (2010 White, 62 Indian)**
Sex, age and survey year	1.06 (0.71, 1.42)***	<0.001 [<0.001 if categorical]	0.63 (0.19, 1.06)***	0.97 (0.26, 1.69)**	2.04 (1.43, 2.64) ***	1.04 (0.66, 1.42)***	0.03 [0.07 if categorical]	0.69 (0.10, 1.28)*	0.94 (0.31, 1.57)**	1.88 (1.12, 2.63)***
Plus academic difficulties and learning difficulties	0.79 (0.43, 1.14)***	<0.001 [0.002 if categorical]	0.35 (-0.12, 0.82)	0.63 (-0.04, 1.30) [p=0.07]	1.72 (1.10, 2.33) ***	0.74 (0.40, 1.07)***	0.13 [0.08 if categorical]	0.47 (-0.09, 1.04) [p=0.10]	0.54 (0.03, 1.06)*	1.54 (0.81, 2.28)***
Plus family type and parental divorce	0.50 (0.16, 0.84)**	0.002 [0.01 if categorical]	0.19 (-0.27, 0.66)	0.32 (-0.34, 0.99)	1.37 (0.77, 1.98) ***	0.49 (0.15, 0.83)**	0.33 [0.12 if categorical]	0.33 (-0.22, 0.89)	0.27 (-0.28, 0.82)	1.19 (0.45, 1.94)**
Plus area, school and family SEP, <u>except</u> parent education	0.57 (0.21, 0.94)**	0.008 [0.03 if categorical]	0.31 (-0.19, 0.81)	0.39 (-0.28, 1.06)	1.19 (0.53, 1.85)***	0.53 (0.16, 0.90)**	0.57 [0.19 if categorical]	0.51 (-0.08, 1.10) [p=.09]	0.24 (-0.35, 0.82)	1.01 (0.15, 1.87)*
Plus other family composition and stress	0.60 (0.24, 0.96)**	0.02 [0.07 if categorical]	0.41 (-0.10, 0.93)	0.41 (-0.28, 1.09)	1.11 (0.43, 1.78)**	0.58 (0.20, 0.96)**	0.73 [0.20 if categorical]	0.53 (-0.07, 1.14) [p=0.09]	0.28 (-0.32, 0.89)	0.99 (0.11, 1.87)*
Plus other child variables	0.51 (0.15, 0.87)**	0.01 [0.05 if categorical]	0.31 (-0.20, 0.81)	0.34 (-0.33, 1.01)	1.00 (0.35, 1.64)**	0.58 (0.20, 0.97)**	0.72 [0.20 if categorical]	0.51 (-0.10, 1.12) [p=0.10]	0.28 (-0.33, 0.89)	1.03 (0.16, 1.91)*
Plus family functioning	0.73 (0.36, 1.09)***	0.007 [0.04 if categorical]	0.45 (-0.05, 0.94) [p=0.08]	0.61 (-0.06, 1.28) [p=0.08]	1.25 (0.62, 1.88)***	0.66 (0.28, 1.04)**	0.69 [0.17 if categorical]	0.55 (-0.05, 1.15) [p=0.07]	0.38 (-0.23, 0.99)	1.16 (0.28, 2.03)*

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Note that data on parent education was missing on 56 individuals for the parent analyses and 25 individuals for the teacher analyses, and these individuals are excluded from these analyses.

Table 14.54: Effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity; stratified analyses by household income

	Parent externalising score					Teacher externalising score				
Adjusted for:	**Full population (13 107 White, 309 Indian)**	**p-value for interaction with household income**	**Income £600 and over (4608 White, 90 Indian)**	**Income £300-599 (4361 White, 99 Indian)**	**Income £0-£299 (4138 White, 120 Indian)**	**Full population (10,251 White, 221 Indian)**	**p-value for interaction with household income**	**Income £600 and over (3742 White, 68 Indian)**	**Income £300-599 (3412 White, 72 Indian)**	**Income £0-£299 (3097 White, 81 Indian)**
Sex, age and survey year	0.87 (0.50, 1.25)***	<0.001 [0.002 if categorical]	0.36 (-0.14, 0.85)	0.48 (-0.19, 1.15)	1.94 (1.32, 2.56)***	0.95 (0.52, 1.38)***	0.06 [<0.001 if categorical]	1.03 (0.45, 1.61)***	0.23 (-0.52, 0.97)	1.91 (1.31, 2.51)***
Plus academic difficulties and learning difficulties	0.61 (0.25, 0.98)**	0.001 [0.006 if categorical]	0.01 (-0.45, 0.48)	0.40 (-0.25, 1.05)	1.47 (0.80, 2.13)***	0.63 (0.27, 0.99)**	0.04 [<0.001 if categorical]	0.56 (0.08, 1.05)*	0.12 (-0.51, 0.76)	1.35 (0.72, 1.98)***
Plus family type and parental divorce	0.32 (-0.03, 0.67) [p=0.07]	0.006 [0.02 if categorical]	-0.09 (-0.55, 0.37)	0.22 (-0.42, 0.86)	0.99 (0.35, 1.64)**	0.38 (0.01, 0.75)*	0.26 [<0.001if categorical]	0.48 (0.00, 0.96)*	-0.08 (-0.71, 0.54)	0.84 (0.17, 1.50)*
Plus area, school and family SEP, <u>except</u> household income	0.40 (0.04, 0.77)*	0.02 [0.04 if categorical]	-0.02 (-0.48, 0.45)	0.34 (-0.34, 1.01)	0.77 (0.11, 1.44)*	0.44 (0.05, 0.83)*	0.54 [<0.001 if categorical]	0.69 (0.22, 1.17)**	-0.13 (-0.79, 0.54)	0.79 (0.08, 1.51)*
Plus other family composition and stress	0.46 (0.09, 0.83)*	0.02 [0.07 if categorical]	0.02 (-0.46, 0.51)	0.49 (-0.24, 1.22)	0.74 (0.09, 1.40)*	0.51 (0.11, 0.91)*	0.71 [0.006 if categorical]	0.75 (0.27, 1.23)**	-0.11 (-0.81, 0.60)	0.81 (0.10, 1.51)*
Plus other child variables	0.37 (0.00, 0.74)*	0.01 [0.05 if categorical]	-0.04 (-0.52, 0.45)	0.34 (-0.39, 1.06)	0.71 (0.06, 1.36)*	0.51 (0.11, 0.91)*	0.75 [0.007 if categorical]	0.74 (0.25, 1.23)**	-0.14 (-0.85, 0.57)	0.81 (0.10, 1.52)*
Plus family functioning	0.60 (0.21, 0.98)**	0.01 [0.02 if categorical]	0.07 (-0.41, 0.56)	0.60 (-0.13, 1.33)	1.00 (0.36, 1.65)**	0.58 (0.18, 0.98)***	0.67 [0.008 if categorical]	0.76 (0.28, 1.25)**	-0.06 (-0.78, 0.65)	0.93 (0.23, 1.64)*

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Note that data on parent education was missing on 813 individuals for the parent analyses and 560 individuals for the teacher analyses, and these individuals are excluded from these analyses.

Table 14.55: Effect of adjustment for variables in all Levels upon the regression coefficient for White (vs. Indian) ethnicity; stratified analyses by tenure

	Parent externalising score				Teacher externalising score			
Adjusted for:	**Full population (13 864 White, 360 Indian)**	**p-value for interaction with tenure**	**Owner occupier (9854 White, 320 Indian)**	**Renting (4010 White, 40 Indian)**	**Full population (10,773 White, 256 Indian)**	**p-value for interaction with tenure**	**Owner occupier (7787 White, 228 Indian)**	**Renting (2986 White, 28 Indian)**
Sex, age and survey year	1.08 (0.73, 1.43)***	0.02	0.56 (0.17, 0.95)**	2.10 (0.96, 3.25)***	1.05 (0.67, 1.43)***	<0.001	0.47 (0.04, 0.89)*	2.81 (2.02, 3.60)***
Plus academic difficulties and learning difficulties	0.81 (0.45, 1.16)***	0.05	0.48 (0.12, 0.85)*	1.68 (0.43, 2.93)**	0.76 (0.43, 1.10)***	<0.001	0.40 (0.05, 0.76)*	2.37 (1.41, 3.34)***
Plus family type and parental divorce	0.52 (0.17, 0.86)**	0.10	0.35 (-0.01, 0.72) [p=0.06]	1.26 (0.02, 2.50)*	0.51 (0.17, 0.85)***	0.001	0.27 (-0.09, 0.62)	2.03 (1.03, 3.02)***
Plus area, school and family SEP, <u>except</u> tenure	0.72 (0.36, 1.08)***	0.13	0.52 (0.13, 0.91)**	1.18 (0.02, 2.34)*	0.67 (0.30, 1.04)***	0.003	0.38 (-0.01, 0.77) [p=0.05]	2.05 (1.04, 3.07)***
Plus other family composition and stress	0.72 (0.36, 1.09)***	0.19	0.58 (0.19, 0.97)**	1.03 (-0.10, 2.17) [p=0.08]	0.71 (0.33, 1.09)***	0.005	0.43 (0.03, 0.83)*	1.98 (0.96, 2.99)***
Plus other child variables	0.62 (0.26, 0.98)**	0.16	0.49 (0.11, 0.88)*	0.88 (-0.23, 1.99)	0.70 (0.32, 1.09)***	0.005	0.43 (0.02, 0.83)*	1.98 (0.97, 2.99)***
Plus family functioning	0.82 (0.46, 1.19)***	0.33	0.69 (0.29, 1.09)**	1.13 (0.10, 2.16)*	0.77 (0.39, 1.16)***	0.008	0.49 (0.09, 0.89)*	2.04 (1.03, 3.05)***

*p<0.05, **p<0.01, ***p<0.001. Table presents the regression coefficients for White (vs. Indian) ethnicity from linear regression. Note that data on tenure was missing on 5 individuals for the parent analyses and 3 individuals for the teacher analyses, and these individuals are excluded from these analyses.

Chapter 15 Appendix 3: Publications arising to date from this PhD

Listed in the order in which they are cited in this PhD:

1. Patel, V. and A. Goodman, Researching protective and promotive factors in mental health. Int J Epidemiol, 2007. 36(4): p. 703-7.
2. Goodman, A., V. Patel, and D.A. Leon, Child mental health differences amongst ethnic groups in Britain: a systematic review. BMC Public Health, 2008. 8: p. 258.
3. Goodman, A. and R. Goodman, Strengths and difficulties questionnaire as a dimensional measure of child mental health. J Am Acad Child Adolesc Psychiatry, 2009. 48(4): p. 400-3.
4. Goodman, A. and R. Gatward, Who are we missing? Area deprivation and survey participation. Eur J Epidemiol, 2008. 23(6): p. 379-87.
5. Heiervang, E., A. Goodman, and R. Goodman, The Nordic advantage in child mental health: separating health differences from reporting style in a cross-cultural comparison of psychopathology. J Child Psychol Psychiatry, 2008. 49(6): p. 678-85.
6. Goodman, A. and T. Ford, Validation of the Ford score as a measure for predicting the level of emotional and behavioural problems in mainstream schools. Research in Education, 2008. 80(1): p. 1-14.

Chapter 16 References

1. Silverman, D., *Doing Qualitative Research: A Practical Handbook*. 2000, London: Sage.
2. Meltzer, H., et al., *Mental health of children and adolescents in Great Britain*. 2000, London: The Stationery Office.
3. Green, H., et al., *Mental health of children and young people in Great Britain, 2004*. 2005: Palgrave MacMillan.
4. Verhulst, F.C. and M. Koot, *Child Psychiatric Epidemiology: Concepts, Methods, and Findings*. 1995, Newbury Park: Sage.
5. WHO, *Caring for children and adolescents with mental disorders*. 2003, Geneva: World Health Organization.
6. Rutter, M. and D.J. Smith, eds. *Psychosocial disorders in young people*. 1995, Wiley: Chichester.
7. Collishaw, S., et al., *Time trends in adolescent mental health.* J Child Psychol Psychiatry, 2004. **45**(8): p. 1350-62.
8. Graham, P., *Mental health must be "centre stage" in child welfare.* Arch Dis Child 2000. **83**(1): p. 4-7.
9. Patel, V. and A. Goodman, *Researching protective and promotive factors in mental health.* Int J Epidemiol, 2007. **36**(4): p. 703-7.
10. Hackett, R., *Gujarati children in Manchester*, in *Ethnicity: An Agenda for Mental Health*, D. Bhugra and V. Bahl, Editors. 1999, Gaskell: London.
11. Maugham, B., *Mental health*, in *The health of our children*, B. Botting, Editor. 1995, OPCS/HMSO London.
12. Ramchandani, P., *The epidemiology of mental health problems in children and adolescents from minority ethnic groups in the UK*, in *Mental Health Services for Minority Ethnic Children and Adolescents*, M. Malek and C. Joughin, Editors. 2004, Jessica Kingsley: London.
13. Vyas, I., *Practice of child psychiatry in multi-ethnic Britain: Some issues and solutions*, in *Ethnicity: An Agenda for Mental Health*, D. Bhugra and V. Bahl, Editors. 1999, Gaskell: London.
14. Littlewood, R. and M. Lipsedge, *Aliens and Alienists: Ethnic minorities and psychiatry*. 3rd ed. 1997, London: Routledge.
15. Mason, D., *Changing ethnic disadvantage: an overview*, in *Explaining Ethnic Differences*, D. Mason, Editor. 2003, The Policy Press: Bristol.
16. WHO, *Mental health: strengthening mental health promotion (Fact sheet no. 220)*. 2007, Geneva: World Health Organisation.
17. Lehtinen, V., et al., *The Intrinsic Value of Mental Health*, in *Promoting mental health: Concepts, Emerging evidence, Practice*, H. Herrman, S. Saxena, and R. Moodie, Editors. 2005, World Health Organisation: Geneva.
18. Thornicroft, G.J., *Shunned: Discrimination Against People with Mental Illness*. 2006, Oxford: Oxford University Press.
19. Herrman, H., et al., *Introduction: Promoting Mental Health as a Public Health Priority*, in *Promoting mental health: Concepts, Emerging evidence, Practice*, H. Herrman, S. Saxena, and R. Moodie, Editors. 2005, World Health Organisation: Geneva.

20. Rose, G. and S. Day, *The population mean predicts the number of deviant individuals.* BMJ, 1990. **301**(6759): p. 1031-4.
21. Veerman, J.L., et al., *Population prevalence of depression and mean Beck Depression Inventory score.* Br J Psychiatry, 2009. **195**(6): p. 516-9.
22. McKenzie, K., *Social risk, mental health and social capital*, in *Social Capital and Mental Health* K. McKenzie and T. Harpham, Editors. 2006, Jessica Kingsley Publishers: London.
23. WHO, *Prevention and promotion in mental health*. 2002, Geneva: World Health Organisation.
24. Tudor, L., *Mental health promotion: paradigms and practice*. 1996, London: Routledge.
25. Huppert, F., N. Baylis, and B. Keverne, *The Science of Well-Being*. 2007, Oxford: Oxford University Press.
26. Zubrick, S. and V. Kovess-Masfety, *Indicators of Mental Health*, in *Promoting mental health : Concepts, Emerging evidence, Practice*, H. Herrman, S. Saxena, and R. Moodie, Editors. 2005, World Health Organisation: Geneva.
27. Huppert, F., *Positive mental health in individuals and populations*, in *The Science of Well-Being*, F. Huppert, N. Baylis, and B. Keverne, Editors. 2007, Oxford University Press: Oxford.
28. Linley, P.A. and S. Joseph, *Positive psychology in practice*. 2004, Hoboken, New Jersey: John Wiley and Sons Inc.
29. Diener. E and R.E. Lucas, *Subjective emotional well-being*, in *Handbook of emotions*, M. Lewis and J.M. Haviland-Jones, Editors. 2000, Guilford press: New York.
30. Keyes, C.L., *The mental health continuum: from languishing to flourishing in life.* J Health Soc Behav, 2002. **43**(2): p. 207-22.
31. Huppert, F.A. and J.E. Whittington, *Evidence for the independence of positive and negative well-being: implications for quality of life assessment.* Br J Health Psychol, 2003. **8**(Pt 1): p. 107-22.
32. Ravens-Sieberer, U., et al., *The KIDSCREEN-52 Quality of Life Measure for Children and Adolescents: Psychometric Results from a Cross-Cultural Survey in 13 European Countries.* Value Health, 2007.
33. UNICEF, *Child poverty in perspective: An overview of child well-being in rich countries, Innocenti Report Card 7*. 2007, Florence: UNICEF Innocenti Research Centre.
34. Goodman, R. and S. Scott, *Child Psychiatry*. 1997, Oxford: Blackwell Scientific Publications.
35. Taylor, E., et al., *The Epidemiology of Childhood Hyperactivity, Maudsley Monograph No. 33*. 1991, Oxford: Oxford University Press.
36. Klein, R.G. and D.S. Pine, *Anxiety Disorders*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
37. Harrington, R., *Affective disorders*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
38. Earls, F. and E. Mezzacappa, *Conduct and oppositional disorders*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
39. Rutter, M., H. Giller, and A. Hagell, *Antisocial behaviour by young people*. 1998, Cambridge: Cambridge University Press.

40. Buitelaar, J.K. and H. Van Engeland, *Epidemiological approaches*, in *Hyperactivity disorders of childhood*, S. Sandberg, Editor. 1996, Cambridge University Press: Cambridge.
41. Schachar, R. and Tannock, *Syndromes of hyperactivity and attention deficit*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
42. Achenbach, T.M. and L.A. Rescorla, *Multicultural understanding of child and adolescent psychopathology : implications for mental health assessment*. 2007, New York: Guilford Press.
43. Angold, A., E.J. Costello, and A. Erkanli, *Comorbidity*. J Child Psychol Psychiatry, 1999. **40**(1): p. 57-87.
44. Patterson, G.R. and M. Stoolmiller, *Replications of a dual failure model for boys' depressed mood.* J Consult Clin. Psychol, 1991. **59**: p. 491-498.
45. Fergusson, D.M., M.T. Lynskey, and L.J. Horwood, *Origins of comorbidity between conduct and affective disorders.* J Am Acad Child Adolesc Psychiatry, 1996. **35**(4): p. 451-60.
46. O'Connor, T.G., et al., *Genetic contributions to continuity, change, and co-occurrence of antisocial and depressive symptoms in adolescence.* J Child Psychol Psychiatry, 1998. **39**(3): p. 323-36.
47. Stringaris, A. and R. Goodman, *Mood lability and psychopathology in youth.* Psychol Med, 2008: p. 1-9.
48. Meltzer, H., et al., *Persistence, onset, risk factors and outcomes of childhood mental disorders*. 2003, London: The Stationery Office.
49. Costello, E.J., H. Egger, and A. Angold, *10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(10): p. 972-86.
50. Willoughby, M.T., *Developmental course of ADHD symptomatology during the transition from childhood to adolescence: a review with recommendations.* J Child Psychol Psychiatry, 2003. **44**(1): p. 88-106.
51. Lara, C., et al., *Childhood predictors of adult attention-deficit/hyperactivity disorder: results from the World Health Organization World Mental Health Survey Initiative.* Biol Psychiatry, 2009. **65**(1): p. 46-54.
52. Patel, V., et al., *Mental health of young people: a global public-health challenge.* Lancet, 2007. **369**(9569): p. 1302-13.
53. Lopez, A.D., et al., *Global Burden of Disease and Risk Factors*. 2006, Oxford: The World Bank and Oxford University Press.
54. Wasserman, D., Q. Cheng, and G.X. Jiang, *Global suicide rates among young people aged 15-19.* World Psychiatry, 2005. **4**(2): p. 114-20.
55. Fergusson, D.M., L.J. Woodward, and L.J. Horwood, *Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood.* Psychol Med, 2000. **30**(1): p. 23-39.
56. Marttunen, M.J., H.M. Aro, and J.K. Lonnqvist, *Adolescence and suicide: a review of psychological autopsy studies.* European Child and Adolescent Psychiatry, 1993. **2**: p. 10-18.
57. Klein, R.G. and S. Mannuzza, *Long-term outcome of hyperactive children: a review*. J Am Acad Child Adolesc Psychiatry, 1991. **30**(3): p. 383-7.

58. Fergusson, D.M. and L.J. Horwood, *The Christchurch Health and Development Study: review of findings on child and adolescent mental health.* Aust N Z J Psychiatry, 2001. **35**(3): p. 287-96.
59. Lynskey, M.T. and D.M. Fergusson, *Childhood conduct problems, attention deficit behaviors, and adolescent alcohol, tobacco, and illicit drug use.* J Abnorm Child Psychol, 1995. **23**(3): p. 281-302.
60. Robins, L.N. and R.K. Price, *Adult disorders predicted by childhood conduct problems: results from the NIMH Epidemiologic Catchment Area project.* Psychiatry, 1991. **54**(2): p. 116-32.
61. Fergusson, D.M., J.M. Boden, and L.J. Horwood, *Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes.* Br J Psychiatry, 2007. **191**: p. 335-42.
62. Fergusson, D.M. and L.J. Woodward, *Mental health, educational, and social role outcomes of adolescents with depression.* Arch Gen Psychiatry, 2002. **59**(3): p. 225-31.
63. Woodward, L.J. and D.M. Fergusson, *Life course outcomes of young people with anxiety disorders in adolescence.* J Am Acad Child Adolesc Psychiatry, 2001. **40**(9): p. 1086-93.
64. Bardone, A.M., et al., *Adult physical health outcomes of adolescent girls with conduct disorder, depression, and anxiety.* J Am Acad Child Adolesc Psychiatry, 1998. **37**(6): p. 594-601.
65. Henderson, M., M. Hotopf, and D.A. Leon, *Childhood temperament and long-term sickness absence in adult life.* British Journal of Psychiatry, 2009. **194**: p. 220-223.
66. Mrazek, D.A., *Psychiatric aspects of somatic disease and disorders*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
67. Goodman, R., *Brain disorders*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
68. Rutter, M., J. Tizard, and K. Whitmore, *Education, Health and Behaviour*. 1970, London: Longmans.
69. Goodman, R. and J. Stevenson, *A twin study of hyperactivity--I. An examination of hyperactivity scores and categories derived from Rutter teacher and parent questionnaires.* J Child Psychol Psychiatry, 1989. **30**(5): p. 671-89.
70. Stevenson, J., et al., *Hyperactivity and spelling disability: testing for shared genetic aetiology.* J Child Psychol Psychiatry, 1993. **34**(7): p. 1137-52.
71. Gillis, J.J., et al., *Attention deficit disorder in reading-disabled twins: evidence for a genetic etiology.* J Abnorm Child Psychol, 1992. **20**(3): p. 303-15.
72. Goodman, R., J. Gledhill, and T. Ford, *Child psychiatric disorder and relative age within school year: cross sectional survey of large population sample.* BMJ, 2003. **327**(7413): p. 472.
73. Emery, R.E. and L. Laumann-Billings, *Child abuse*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
74. Rushton, A. and H. Minnis, *Residential and foster family care*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
75. Meltzer, H., et al., *The mental health of young people looked after by local authorities in England*. 2003, London: The Stationary Office.
76. Graham, P., *The end of adolescence*. 2004, Oxford: Oxford University Press.

77. Goodman, A., et al., *Child, family, school and community risk factors for poor mental health in Brazilian schoolchildren.* J Am Acad Child Adolesc Psychiatry, 2007. **46**(4): p. 448-56.
78. Patel, V., et al., *Social determinants of mental disorders*, in *Priority Public Health Conditions: From Learning to Action on Social Determinants of Health (forthcoming)*, E. Blas and S.K. A, Editors. In press, World Health Organization: Geneva.
79. Alyahri, A. and R. Goodman, *Harsh corporal punishment of Yemeni children: occurrence, type and associations.* Child Abuse Negl, 2008. **32**(8): p. 766-73.
80. Scott, S., et al., *Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice.* BMJ, 2001. **323**(7306): p. 194-8.
81. DiScala, C., et al., *Injuries to children with attention deficit hyperactivity disorder.* Pediatrics, 1998. **102**(6): p. 1415-21.
82. Leibson, C.L., et al., *Use and costs of medical care for children and adolescents with and without attention-deficit/hyperactivity disorder.* JAMA, 2001. **285**(1): p. 60-6.
83. Brehaut, J.C., et al., *Childhood behavior disorders and injuries among children and youth: a population-based study.* Pediatrics, 2003. **111**(2): p. 262-9.
84. Pastor, P.N. and C.A. Reuben, *Identified attention-deficit/hyperactivity disorder and medically attended, nonfatal injuries: US school-age children, 1997-2002.* Ambul Pediatr, 2006. **6**(1): p. 38-44.
85. Levin, H., et al., *Symptoms of attention-deficit/hyperactivity disorder following traumatic brain injury in children.* J Dev Behav Pediatr, 2007. **28**(2): p. 108-18.
86. Keenan, H.T., G.C. Hall, and S.W. Marshall, *Early head injury and attention deficit hyperactivity disorder: retrospective cohort study.* BMJ, 2008. **337**: p. a1984.
87. Sandberg, S. and M. Rutter, *The role of acute life stressors,* in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
88. Black, D., *Bereavement,* in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
89. Goodyer, I., C. Wright, and P. Altham, *The friendships and recent life events of anxious and depressed school-age children.* Br J Psychiatry, 1990. **156**: p. 689-98.
90. Goodyer, I.M., *Developmental psychopathology: the impact of recent life events in anxious and depressed school-age children.* J R Soc Med, 1994. **87**(6): p. 327-9.
91. Goodyer, I.M., C. Wright, and P.M. Altham, *Recent friendships in anxious and depressed school age children.* Psychol Med, 1989. **19**(1): p. 165-74.
92. Goodyer, I.M., et al., *Short-term outcome of major depression: II. Life events, family dysfunction, and friendship difficulties as predictors of persistent disorder.* J Am Acad Child Adolesc Psychiatry, 1997. **36**(4): p. 474-80.
93. Brown, R.A., et al., *Cigarette smoking, major depression, and other psychiatric disorders among adolescents.* J Am Acad Child Adolesc Psychiatry, 1996. **35**(12): p. 1602-10.
94. Breslau, N., *Psychiatric comorbidity of smoking and nicotine dependence.* Behav Genet, 1995. **25**(2): p. 95-101.
95. Burke, J.D., et al., *Inattention as a key predictor of tobacco use in adolescence.* J Abnorm Psychol, 2007. **116**(2): p. 249-59.
96. Costello, E.J., et al., *Development of psychiatric comorbidity with substance abuse in adolescents: effects of timing and sex.* J Clin Child Psychol, 1999. **28**(3): p. 298-311.

97. Fergusson, D.M., L.J. Horwood, and E.M. Ridder, *Conduct and attentional problems in childhood and adolescence and later substance use, abuse and dependence: results of a 25-year longitudinal study*. Drug Alcohol Depend, 2007. **88 Suppl 1**: p. S14-26.
98. Griesler, P.C., et al., *Comorbidity of psychiatric disorders and nicotine dependence among adolescents: findings from a prospective, longitudinal study*. J Am Acad Child Adolesc Psychiatry, 2008. **47**(11): p. 1340-50.
99. Storr, C.L., B.A. Reboussin, and J.C. Anthony, *Early childhood misbehavior and the estimated risk of becoming tobacco-dependent.* Am J Epidemiol, 2004. **160**(2): p. 126-30.
100. White, H.R., et al., *Psychopathology as a predictor of adolescent drug use trajectories.* Psychol Addict Behav, 2001. **15**(3): p. 210-8.
101. Wittchen, H.-U., et al., *Cannabis use and cannabis use disorders and their relationship to mental disorders: A 10-year prospective-longitudinal community study in adolescents.* Drug and Alcohol Dependence, 2007. **88**(Supplement 1): p. S60-S70.
102. Boyle, M.H., et al., *Predicting substance use in late adolescence: results from the Ontario Child Health Study follow-up.* Am J Psychiatry, 1992. **149**(6): p. 761-7.
103. King, S.M., W.G. Iacono, and M. McGue, *Childhood externalizing and internalizing psychopathology in the prediction of early substance use.* Addiction, 2004. **99**(12): p. 1548-59.
104. Sung, M., et al., *Effects of age at first substance use and psychiatric comorbidity on the development of substance use disorders.* Drug Alcohol Depend, 2004. **75**(3): p. 287-99.
105. Pedersen, W., A. Mastekaasa, and L. Wichstrom, *Conduct problems and early cannabis initiation: a longitudinal study of gender differences.* Addiction, 2001. **96**(3): p. 415-31.
106. Cohen, P., et al., *Personality disorders in early adolescence and the development of later substance use disorders in the general population.* Drug and Alcohol Dependence, 2007. **88**(Supplement 1): p. S71-S84.
107. Pardini, D., H.R. White, and M. Stouthamer-Loeber, *Early adolescent psychopathology as a predictor of alcohol use disorders by young adulthood.* Drug Alcohol Depend, 2007. **88 Suppl 1**: p. S38-49.
108. Costello, E.J., et al., *Pubertal maturation and the development of alcohol use and abuse*. Drug and Alcohol Dependence, 2007. **88**(Supplement 1): p. S50-S59.
109. Henry, B., et al., *The importance of conduct problems and depressive symptoms in predicting adolescent substance use.* J Abnorm Child Psychol, 1993. **21**(5): p. 469-80.
110. Flory, K. and D.R. Lynam, *The relation between attention deficit hyperactivity disorder and substance abuse: what role does conduct disorder play?* Clin Child Fam Psychol Rev, 2003. **6**(1): p. 1-16.
111. Wang, M.Q., et al., *Prospective social-psychological factors of adolescent smoking progression.* J Adolesc Health, 1999. **24**(1): p. 2-9.
112. Patton, G.C., et al., *Depression, anxiety, and smoking initiation: a prospective study over 3 years.* Am J Public Health, 1998. **88**(10): p. 1518-22.
113. Sonntag, H., et al., *Are social fears and DSM-IV social anxiety disorder associated with smoking and nicotine dependence in adolescents and young adults?* Eur Psychiatry, 2000. **15**(1): p. 67-74.

114. Hofler, M., et al., *Covariates of cannabis use progression in a representative population sample of adolescents: a prospective examination of vulnerability and risk factors.* Addiction, 1999. **94**(11): p. 1679-94.
115. Dierker, L.C., et al., *Testing the dual pathway hypothesis to substance use in adolescence and young adulthood.* Drug Alcohol Depend, 2007. **87**(1): p. 83-93.
116. Isensee, B., et al., *Smoking increases the risk of panic: findings from a prospective community study.* Arch Gen Psychiatry, 2003. **60**(7): p. 692-700.
117. Achenbach, T.M., et al., *Six-year predictors of problems in a national sample: IV. Young adult signs of disturbance.* J Am Acad Child Adolesc Psychiatry, 1998. **37**(7): p. 718-27.
118. Patton, G.C., et al., *Cannabis use and mental health in young people: cohort study.* Bmj, 2002. **325**(7374): p. 1195-8.
119. Brook, J.S., P. Cohen, and D.W. Brook, *Longitudinal study of co-occurring psychiatric disorders and substance use.* J Am Acad Child Adolesc Psychiatry, 1998. **37**(3): p. 322-30.
120. Breslau, N. and D.F. Klein, *Smoking and panic attacks: an epidemiologic investigation.* Arch Gen Psychiatry, 1999. **56**(12): p. 1141-7.
121. Johnson, J.G., et al., *Association between cigarette smoking and anxiety disorders during adolescence and early adulthood.* Jama, 2000. **284**(18): p. 2348-51.
122. Breslau, N., S.P. Novak, and R.C. Kessler, *Daily smoking and the subsequent onset of psychiatric disorders.* Psychol Med, 2004. **34**(2): p. 323-33.
123. Kandel, D.B., et al., *The consequences in young adulthood of adolescent drug involvement. An overview.* Arch Gen Psychiatry, 1986. **43**(8): p. 746-54.
124. Breslau, N., et al., *Major depression and stages of smoking. A longitudinal investigation.* Arch Gen Psychiatry, 1998. **55**(2): p. 161-6.
125. Arseneault, L., et al., *Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study.* Bmj, 2002. **325**(7374): p. 1212-3.
126. Brook, J.S., E. Schuster, and C. Zhang, *Cigarette smoking and depressive symptoms: a longitudinal study of adolescents and young adults.* Psychol Rep, 2004. **95**(1): p. 159-66.
127. McGee, R., et al., *A longitudinal study of cannabis use and mental health from adolescence to early adulthood.* Addiction, 2000. **95**(4): p. 491-503.
128. Clark, C., et al., *Psychological symptoms and physical health and health behaviours in adolescents: a prospective 2-year study in East London.* Addiction, 2007. **102**(1): p. 126-35.
129. Silberg, J., et al., *Genetic and environmental risk factors in adolescent substance use.* J Child Psychol Psychiatry, 2003. **44**(5): p. 664-76.
130. McMunn, A.M., et al., *Children's emotional and behavioural well-being and the family environment: findings from the Health Survey for England.* Soc Sci Med, 2001. **53**(4): p. 423-40.
131. Kiernan, K.E., *The impact of family disruption in childhood on transition made in young adult life.* Population Studies, 1992. **46**: p. 213-234.
132. Maughan, B. and A. Pickles, *Adopted and illegitimate children growing up*, in *Straight and devious pathways from childhood to adulthood*, L. Robins and M. Rutter, Editors. 1990, Cambridge University Press: Cambridge.
133. Wadsworth, J., et al., *The influence of family type on children's behaviour and development at five years.* J Child Psychol Psychiatry, 1985. **26**(2): p. 245-54.

134. Reid, W.J. and A. Crisafulli, *Marital discord and child behavior problems: a meta-analysis.* J Abnorm Child Psychol, 1990. **18**(1): p. 105-17.
135. Depner, C.E., E.V. Leino, and A. Chun, *Interparental conflict and child adjustment: A decade review and meta-analysis.* Family & Conciliation Courts Review, 1992. **30**(3): p. 323-341.
136. Davies, P.T. and M. Windle, *Gender-specific pathways between maternal depressive symptoms, family discord, and adolescent adjustment.* Dev Psychol, 1997. **33**(4): p. 657-68.
137. Lorenz, G., et al., *Marital discord and psychiatric disorder in children and adolescents* Journal of Child and Family Studies, 1995. **4**(3): p. 341-358.
138. Patterson, G.R., et al., *The aggressive child: victim and architect of a coercive system. Behaviour modifications and families: theory and research*. 1976, New York: Brunner/Mazel.
139. Friedman, R.J. and P. Lindsay Chase-Lansdale, *Chronic adversities*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
140. Field, T., *Infants of depressed mothers.* Infant behaviour and development, 1995. **18**: p. 1-13.
141. Leadbeater, B.J., S.J. Bishop, and C.C. Raver, *Quality of mother-toddler interactions, maternal depressive symptoms and behaviour problems in preschoolers of adolescent mothers.* Dev Psychol, 1996. **32**: p. 280-288.
142. Rubin, K.H., et al., *Dyadic play behaviors of children of well and depressed mothers.* Dev Psychol, 1991. **3**: p. 243-152.
143. Goodman, S.H., et al., *Social and emotional competence in children of depressed mothers.* Child Dev, 1993. **64**(2): p. 516-31.
144. Ramchandani, P., et al., *Paternal depression in the postnatal period and child development: a prospective population study.* Lancet, 2005. **365**(9478): p. 2201-5.
145. Ramchandani, P.G., et al., *Depression in men in the postnatal period and later child psychopathology: a population cohort study.* J Am Acad Child Adolesc Psychiatry, 2008. **47**(4): p. 390-8.
146. Silberg, J., et al., *Genetic moderation of environmental risk for depression and anxiety in adolescent girls.* Br J Psychiatry, 2001. **179**: p. 116-21.
147. Hammen, C., D. Burge, and C. Adrian, *Timing of mother and child depression in a longitudinal study of children at risk.* J Consult Clin Psychol, 1991. **59**(2): p. 341-5.
148. Rutter, M., *Afterword. Maternal depression and infant development: cause and consequences; sensitivity and specificity*, in *Postpartum depression and child development*, L. Murray and P. Cooper, Editors. 1997, Guilford: New York.
149. Duncan, G.J. and J. Brooks-Gunn, *Consequences of growing up poor*. 1997, New York: Russell Sage Foundation.
150. Garrett, P., N. Ng'andu, and J. Ferron, *Poverty experiences of young children and the quality of their home environments.* Child Dev, 1994. **65**(2 Spec No): p. 331-45.
151. Kalff, A.C., et al., *Factors affecting the relation between parental education as well as occupation and problem behaviour in Dutch 5- to 6-year-old children.* Soc Psychiatry Psychiatr Epidemiol, 2001. **36**(7): p. 324-31.
152. Kalff, A.C., et al., *Neighbourhood level and individual level SES effects on child problem behaviour: a multilevel analysis.* J Epidemiol Community Health, 2001. **55**(4): p. 246-50.

153. Ecob, R. and R. Williams, *Sampling Asian minorities to assess health and welfare.* J Epidemiol Community Health, 1991. **45**(2): p. 93-101.
154. Halpern, D. and J. Nazroo, *The ethnic density effect: Results from a national community survey of England and Wales.* International Journal of Social Psychiatry, 2000. **46**(1): p. 34-46.
155. Neeleman, J., C. Wilson-Jones, and S. Wessely, *Ethnic density and deliberate self harm; a small area study in South East London.* J Epidemiol Community Health, 2001. **55**(2): p. 85-90.
156. Karlsen, S., J.Y. Nazroo, and R. Stephenson, *Ethnicity, environment and health: putting ethnic inequalities in health in their place.* Soc Sci Med, 2002. **55**(9): p. 1647-61.
157. Fagg, J., et al., *Psychological distress among adolescents, and its relationship to individual, family and area characteristics in East London.* Soc Sci Med, 2006. **63**(3): p. 636-48.
158. Halpern, D., *Minorities and mental health.* Soc Sci Med, 1993. **36**(5): p. 597-607.
159. Rutter, M., *The interplay of nature, nurture, and developmental influences: the challenge ahead for mental health.* Arch Gen Psychiatry, 2002. **59**(11): p. 996-1000.
160. Thapar, A., et al., *Genetic basis of attention deficit and hyperactivity* Br J Psychiatry, 1999. **174**: p. 105-111.
161. Silberg, J., et al., *The influence of genetic factors and life stress on depression among adolescent girls.* Arch Gen Psychiatry, 1999. **56**(3): p. 225-32.
162. Jockin, V., M. McGue, and D.T. Lykken, *Personality and divorce: a genetic analysis.* J Pers Soc Psychol, 1996. **71**(2): p. 288-99.
163. Plomin, R., *Genetics and experience: the interplay between nature and nurture*. 1994, Thousand Oaks, CA: Sage Publications.
164. Thapar, A., et al., *Do intrauterine or genetic influences explain the foetal origins of chronic disease? A novel experimental method for disentangling effects.* BMC Med Res Methodol, 2007. **7**: p. 25.
165. Kim-Cohen, J., et al., *MAOA, maltreatment, and gene-environment interaction predicting children's mental health: new evidence and a meta-analysis.* Mol Psychiatry, 2006. **11**(10): p. 903-13.
166. Rutter, M., et al., *A tri-axial classification of mental disorders in childhood. An international study.* J Child Psychol Psychiatry, 1969. **10**(1): p. 41-61.
167. Taylor, E. and M. Rutter, *Classification: Conceptual Issues and Substantive Findings*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
168. Bird, H.R., et al., *Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico. The use of combined measures.* Arch Gen Psychiatry, 1988. **45**(12): p. 1120-1126.
169. Simonoff, E., et al., *The Virginia Twin Study of Adolescent Behavioral Development. Influences of age, sex, and impairment on rates of disorder.* Arch Gen Psychiatry, 1997. **54**(9): p. 801-8.
170. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).* 1994, Washington, DC: American Psychiatric Association.
171. WHO, *The ICD-10 classification of mental and behavioural disorders; diagnostic criteria for research.* 1993, Geneva: World Health Organisation.

172. Goodman, R., et al., *Rating child psychiatric caseness from detailed case histories.* J Child Psychol Psychiatry, 1996. **37**(4): p. 369-79.
173. Goldberg, D. and P. Huxley, *Common mental disorders: a bio-social model*. 1992, London: Tavistock.
174. Rutter, M. and A. Nikapota, *Culture, Ethnicity, Society and Psychopathology*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
175. Alegria, M. and T. McGuire, *Rethinking a universal framework in the psychiatric symptom-disorder relationship.* J Health Soc Behav, 2003. **44**(3): p. 257-74.
176. Wakefield, J.C., *DSM-IV: Are We Making Diagnostic Progress?* Contemporary Psychology, 1996. **41**(7): p. 646-52.
177. Patel, V. and M. Winston, *'Universality of mental illness' revisited: assumptions, artefacts and new directions.* Br J Psychiatry, 1994. **165**: p. 437-440.
178. Mohler, B., *Cross-cultural issues in research on child mental health.* Child Adolesc Psychiatr Clin N Am, 2001. **10**(4): p. 763-76, ix.
179. Odenwald, M., M. van Duijl, and T. Schmitt, *Psychopathology and culture: disorders of possession and dissociation in intercultural clinical practice*, in *Culture and mental health: a comprehensive textbook*, K. Bhui and D. Bhugra, Editors. 2007, Hodder Arnold: London.
180. Alarcón, R., et al., *Beyond the funhouse mirrors: research agenda on culture and psychiatric diagnosis*, in *A research agenda for DSM-IV*, D. Kupfer, M. First, and D. Regier, Editors. 2002, American Psychiatric Association: Washington DC.
181. Kleinman, A., *Depression, somatisation and the new 'cross-cultural psychiatry'.* Social Science and Medicine, 1977. **11**: p. 3-10.
182. Littlewood, R., *From categories to contexts: a decade of the 'new cross-cultural psychiatry'.* Br J Psychiatry, 1990. **156**: p. 308-27.
183. Kleinman, A., *Anthropology and psychiatry. The role of culture in cross-cultural research on illness.* Br J Psychiatry, 1987. **151**: p. 447-54.
184. Marsella, A., J., et al., *Cross-cultural studies of depressive disorders: An overview*, in *Culture and depression*, A. Kleinman and G. Byron, Editors. 1985, University of California press: Berkeley, CA.
185. Bass, J., et al., *Group interpersonal psychotherapy for depression in rural Uganda: 6-month outcomes: Randomised controlled trial.* The British Journal of Psychiatry, 2006. **188**(6): p. 567-573.
186. Hewitt, L.E. and R.J. Jenkins, *Fundamental Patterns of Maladjustment: the Dynamics of Their Origin*. 1946, Illinois: Michigan Child Guidance Institute.
187. Achenbach, T.M. and C.D. Edelbrok, *Manual for the Child Behaviour Checklist and Revised Child Behaviour Profile*. 1983, Burlington, Vermont: University of Vermont, Department of Psychiatry.
188. Achenbach, T.M., *Manual for the Teacher's Report Form and 1991 Profile*. 1991, Burlington, VT: University of Vermont, Department of Psychiatry.
189. Achenbach, T.M., *Manual for the Youth Self-Report and 1991 Profile*. 1991, Burlington, VT: University of Vermont, Department of Psychiatry.
190. Achenbach, T.M., *Manual for the Child Behavior Checklist 4-18 and 1991 Profile*. 1991, Burlington, VT: University of Vermont, Department of Psychiatry.
191. Achenbach, T.M., et al., *Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings,*

applications, and future directions. J Child Psychol Psychiatry, 2008. **49**(3): p. 251-75.
192. Achenbach, T.M., *The classification of children's psychiatric symptoms: a factor-analytic study.* Psychol Monogr, 1966. **80**(7): p. 1-37.
193. Elander, J. and R. M., *Use and development of the Rutter parents' and teachers' scales.* International Journal of Methods in Psychiatric Research, 1996. **6**(2): p. 63-78.
194. Goodman, R., *The Strengths and Difficulties Questionnaire: a research note.* J Child Psychol Psychiatry, 1997. **38**(5): p. 581-6.
195. Lahey, B.B., et al., *Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology.* J Abnorm Child Psychol, 2008. **36**(2): p. 187-206.
196. Hartman, C.A., et al., *DSM-IV internal construct validity: when a taxonomy meets data.* J Child Psychol Psychiatry, 2001. **42**(6): p. 817-36.
197. Ivanova, M.Y., et al., *The generalizability of the Youth Self-Report syndrome structure in 23 societies.* J Consult Clin Psychol, 2007. **75**(5): p. 729-38.
198. Ivanova, M.Y., et al., *Testing the 8-syndrome structure of the child behavior checklist in 30 societies.* J Clin Child Adolesc Psychol, 2007. **36**(3): p. 405-17.
199. Ivanova, M.Y., et al., *Testing the Teacher's Report Form syndromes in 20 societies.* School Psychology Review, 2007. **36**: p. 468-483.
200. Woerner, W., et al., *The Strengths and Difficulties Questionnaire overseas: evaluations and applications of the SDQ beyond Europe.* Eur Child Adolesc Psychiatry, 2004. **13 Suppl 2**: p. II/47-54.
201. Bird, H., *Epidemiology of childhood disorders in a cross-cultural context.* Journal of Child Psychology and Psychiatry, 1996. **37**(1): p. 35-49.
202. Schneider, B.H., *Cross-cultural comparison as doorkeeper in research on the social and emotional adjustment of children and adolescents.* Dev Psychol, 1998. **34**(4): p. 793-7.
203. Luk, S., *Cross cultural/ethnical aspects*, in *Hyperactivity disorders of Childhood*, S. Sandberg, Editor. 1996, Cambridge UP: Cambridge.
204. Crijnen, A.A.M., T.M. Achenbach, and F.C. Verhulst, *Problems Reported by Parents of Children in Multiple Cultures: The Child Behavior Checklist Syndrome Constructs.* Am J Psychiatry, 1999. **156**(4): p. 569-574.
205. Nikapota, A.D., *Child psychiatry in developing countries.* Br J Psychiatry, 1991. **158**: p. 743-51.
206. de Jong, T.J. and M. Van Ommeren, *Toward a Culture-informed Epidemiology: Combining Qualitative and Quantitative Research in Transcultural Contexts.* Transcultural psychiatry, 2002. **39**(4): p. 422-433.
207. Good, B. and J. Good, *Towards a meaning-centred analysis of popular illness categories: 'fright-illness' and 'heat distress' in Iran*, in *Cultural Conceptions of Mental Health and Therapy*, M.A. J. and W.G. M., Editors. 1982, Reidel: Dordrecht.
208. Robertson, B.A. and A. Kottler, *Cultural issues in the psychiatric assessment of Xhosa children and adolescents.* S Afr Med J, 1993. **83**(3): p. 207-8.
209. Ensink, K. and B. Robertson, *Indigenous Categories of Distress and Dysfunction in South African Xhosa Children and Adolescents as Described by Indigenous Healers* Transcultural Psychiatry, 1996. **137-172**.

210. O’Conner, W. and J. Nazroo, *Ethnic differences in the context and experience of psychiatric illness: a qualitative study*. 2002, London: National Centre for Social Research and Department of Epidemiology and Public Health, Royal Free and University College Medical School.
211. Fenton, F.R. and A. Sadiq-Sangster, *Culture, relativism and the expression of mental distress; South Asian women in Britain.* Sociology of health and illness, 1996. **18**(1): p. 66-85.
212. Krause, I.B., *Sinking heart: a Punjabi communication of distress.* Soc Sci Med, 1989. **29**(4): p. 563-75.
213. Armstrong, C., M. Hill, and J. Secker, *Young people’s perceptions of mental health.* Children and Society, 2000. **14**: p. 60-72.
214. Nikapota, A., et al., *Development of Culturally Appropriate Child Mental Health Services: Perceptions and Use of Services*. 1998, London: Unpublished report to the Department of Health.
215. Hackett, R. and L. Hackett, *Child psychiatry across cultures.* International review of Psychiatry, 1999. **11**: p. 225-235.
216. Remschmidt, H. and M. Belfer, *Mental health care for children and adolescents worldwide: a review.* World Psychiatry, 2005. **4**(3): p. 147-153.
217. Hermans, H.J.M. and H.J.G. Kempen, *Moving Cultures: The Perilous Problems of Cultural Dichotomies in a Globalizing Society.* 53, 1998. **10**(1111-1120).
218. Sturgeon, S. and J. Orley, *Concepts of Mental Health Across the World*, in *Promoting mental health : Concepts, Emerging evidence, Practice*, H. Herrman, S. Saxena, and R. Moodie, Editors. 2005, World Health Organisation: Geneva.
219. Eisenberg, L., *Mindlessness and brainlessness in psychiatry.* Br J Psychiatry, 1986. **148**: p. 497-508.
220. Midgley, M., *Science and poetry*. 2006, London: Routledge.
221. Simpson, L. and B. Akinwale (2006) *Quantifying stability and change in ethnic group*. CCSR Working Paper 2006-05.
222. Plewis, I. and S. Ketende, *Millennium Cohort Survey: Technical report on Response*. 2006, Centre for Longitudinal Studies, Institute of Education: London.
223. Weisz, J.R. and K.L. Eastman, *Cross-national research on child and adolescent psychopathology*, in *The epidemiology of child and adolescent psychopathology.*, F.C. Verhulst and H.M. Koot, Editors. 1995, Oxford University Press: Oxford.
224. van der Vijver, F.J.R. and Y.H. Poortinga, *Towards an integrated analysis of bias in cross-cultural assessment.* Eur J of Psychological Assessment, 1997(13): p. 29-37.
225. Hewitt, J.K., et al., *Genetics and developmental psychopathology: 1. Phenotypic assessment in the Virginia Twin Study of Adolescent Behavioral Development.* J Child Psychol Psychiatry, 1997. **38**(8): p. 943-63.
226. Achenbach, T.M., S.H. McConaughy, and C.T. Howell, *Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity.* Psychol Bull, 1987. **101**(2): p. 213-32.
227. Angold, A., *Diagnostic Interviews with Parents and Children*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
228. Chilcoat, H.D. and N. Breslau, *Does psychiatric history bias mothers' reports? An application of a new analytic approach.* J Am Acad Child Adolesc Psychiatry, 1997. **36**(7): p. 971-9.

229. Borge, A., S. Samuelsen, and M. Rutter, *Observer variance within families: confluence among maternal, paternal and child ratings.* International Journal of Methods in Psychiatric Research, 2001. **10**(1): p. 11-21.
230. Heiervang, E., A. Goodman, and R. Goodman, *The Nordic advantage in child mental health: separating health differences from reporting style in a cross-cultural comparison of psychopathology*. J Child Psychol Psychiatry, 2008. **49**(6): p. 678-85.
231. Mann, E.M., et al., *Cross-cultural differences in rating hyperactive-disruptive behaviors in children.* Am J Psychiatry, 1992. **149**(11): p. 1539-42.
232. Sonuga-Barke, E., et al., *Inter-ethnic bias in teachers' ratings of childhood hyperactivity*. British Journal of Developmental Psychology 1993. **11**(2): p. 187-200.
233. Breslau, N., *Inquiring about the bizarre: false positives in Diagnostic Interview Schedule for Children (DISC) ascertainment of obsessions, compulsions, and psychotic symptoms.* J Am Acad Child Adolesc Psychiatry, 1987. **26**(5): p. 639-44.
234. Brugha, T.S., P.E. Bebbington, and R. Jenkins, *A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population.* Psychol Med, 1999. **29**(5): p. 1013-20.
235. Angold, A, *Structured assessments of psychopathology in children and adolescents*, in *The Instruments of Psychiatric Research*, C. Thompson, Editor. 1989, John Wiley & Sons Ltd: New York.
236. Loeber, R., et al., *Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior.* J Abnorm Child Psychol, 1991. **19**(1): p. 75-95.
237. Blatchford, P., *Pupils' self assessments of academic attainment at 7, 11 and 16 years: effects of sex and ethnic group.* Br J Educ Psychol, 1997. **67 (Pt 2)**: p. 169-84.
238. Rutter, M., et al., *Children of West Indian immigrants. I. Rates of behavioural deviance and of psychiatric disorder.* J Child Psychol Psychiatry, 1974. **15**(4): p. 241-62.
239. Fombonne, E., *Case identification in an epidemiological context*, in *Child and Adolescent Psychiatry*, M. Rutter and E.A. Taylor, Editors. 2002, Blackwell: Oxford.
240. Bird, H.R., M.S. Gould, and B. Staghezza, *Aggregating data from multiple informants in child psychiatry epidemiological research.* J Am Acad Child Adolesc Psychiatry, 1992. **31**: p. 78-85.
241. Modood, T., et al., *Ethnic Minorities in Britain: Diversity and Disadvantage*. 1997, London: Policy Studies Institute.
242. Panico, L., et al., *Ethnic variation in childhood asthma and wheezing illnesses: findings from the Millennium Cohort Study.* Int J Epidemiol, 2007. **36**(5): p. 1093-102.
243. Nazroo, J.Y., *Ethnicity and mental health: findings from a National Community Survey*. 1997, London: Policy Studies Institute.
244. Weich, S., et al., *Common mental disorders and ethnicity in England: the EMPIRIC study.* Psychol Med, 2004. **34**(8): p. 1543-51.
245. Cronbach, L.E. and P.E. Meehl, *Construct validity in psychological tests.* Psychological Bulletin, 1955. **52**: p. 281-302.

246. Ahmad, W.I.U., *Making black people sick: 'race', ideology and health research*, in *'Race' and health in contemporary Britain*, W.I.U. Ahmad, Editor. 1993, Open University Press: Buckingham.
247. Bhopal, R., *Is research into ethnicity and health racist, unsound, or important science?* Bmj, 1997. **314**(7096): p. 1751-6.
248. Senior, P.A. and R. Bhopal, *Ethnicity as a variable in epidemiological research.* Bmj, 1994. **309**(6950): p. 327-30.
249. Nazroo, J.Y., *Genetic, cultural or socio-economic vulnerability? Explaining ethnic inequalities in health.* Sociology of health and illness, 1998. **20**(5): p. 710-730.
250. Davidson, K.W., et al., *Gender as a health determinant and implications for health education.* Health Educ Behav, 2006. **33**(6): p. 731-43; discussion 744-6.
251. Smedley, A., *Race in North America: origin and evolution of a worldview*. 2nd ed. 1999, Boulder Westview Press.
252. Barkin, E., *The Retreat of Scientific Racism: Changing Concepts of Race in Britain and the United States between the World Wars*. 1993, Cambridge: Cambridge UP.
253. Barbujani, G., et al., *An apportionment of human DNA diversity.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4516-9.
254. Drake, A.J. and B.R. Walker, *The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and cardiovascular risk.* J Endocrinol, 2004. **180**(1): p. 1-16.
255. Nature Genetics Supplement, *Genetics for the Human Race* 2004. **36**(11s).
256. Yajnik, C.S., *Early life origins of insulin resistance and type 2 diabetes in India and other Asian countries.* J Nutr, 2004. **134**(1): p. 205-10.
257. Weber, M., *Ethnic groups*, in *Economy and Society, Vol 1 part II*, G. Roth and C. Wittich, Editors. 1978 [1922], University of California Press: Berkeley. p. 385-398.
258. Modood, T., S. Beishon, and S. Virdee, *Changing ethnic identities*. 1994, London: Policy Studies Institute.
259. Ghuman, P. and A. Singh, *Ethnic identity and acculturation of South Asian adolescents: A British perspective.* International Journal of Adolescence and Youth, 1998. **7**(3): p. 227-247.
260. Karlsen, S., *'Black like Beckham'? Moving beyond definitions of ethnicity based on skin colour and ancestry.* Ethn Health, 2004. **9**(2): p. 107-37.
261. White, A., *Social Focus in Brief: Ethnicity 2002*. 2002, Office of National Statistics: London.
262. Modood, T., *Ethnicity and intergenerational identities and adaptations in Britain: the socio-political context.*, in *Ethnicity and Causal Mechanisms*, M. Rutter and M. Tienda, Editors. 2005, Cambridge University Press: Cambridge.
263. Ahmad, W.I.U. and H. Bradby, *Locating ethnicity and health: exploring concepts and contexts.* Sociology of Health & Illness, 2007. **29**(6): p. 795-810.
264. Kendler, K.S., *"A gene for...": the nature of gene action in psychiatric disorders.* Am J Psychiatry, 2005. **162**(7): p. 1243-52.
265. Serre, D. and S. Paabo, *Evidence for gradients of human genetic diversity within and among continents.* Genome Res, 2004. **14**(9): p. 1679-85.
266. Bauchet, M., et al., *Measuring European population stratification with microarray genotype data.* Am J Hum Genet, 2007. **80**(5): p. 948-56.
267. Edwards, A.W., *Human genetic diversity: Lewontin's fallacy.* Bioessays, 2003. **25**(8): p. 798-801.

268. Barnett, A.H., et al., *Type 2 diabetes and cardiovascular risk in the UK south Asian community*. Diabetologia, 2006. **49**(10): p. 2234-46.
269. Stutzmann, F., et al., *Common genetic variation near MC4R is associated with eating behaviour patterns in European populations.* Int J Obes (Lond), 2009.
270. Dhandapany, P.S., et al., *A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia.* Nat Genet, 2009.
271. Hulshoff Pol, H.E., et al., *Prenatal exposure to famine and brain morphology in schizophrenia.* Am J Psychiatry, 2000. **157**(7): p. 1170-2.
272. Brown, A.S., et al., *Further evidence of relation between prenatal famine and major affective disorder.* Am J Psychiatry, 2000. **157**(2): p. 190-5.
273. Hoek, H.W., A.S. Brown, and E. Susser, *The Dutch famine and schizophrenia spectrum disorders.* Soc Psychiatry Psychiatr Epidemiol, 1998. **33**(8): p. 373-9.
274. Krieger, N., et al., *Racism, sexism, and social class: implications for studies of health, disease, and well-being.* Am J Prev Med, 1993. **9**(6 Suppl): p. 82-122.
275. Duster, T., *Medicalisation of race.* The Lancet, 2007. **369**(9562): p. 702-704.
276. Farmer, P., *Infections and inequalities : the modern plagues*. 1999, Berkeley ; London: University of California Press.
277. Scheper-Hughes, N., *Death without weeping: the violence of everyday life in Brazil*. 1992, Berkeley/Oxford: University of California Press.
278. Oliver, M.N. and C. Muntaner, *Researching health inequities among African Americans: the imperative to understand social class.* Int J Health Serv, 2005. **35**(3): p. 485-98.
279. Cochrane, R. and S.P. Sashidharan, *Mental health and ethnic minorities: a review of the literature and implications for services*, in *Ethnicity and Health*, W.I.U. Ahmad, T.A. Sheldon, and O. Stuart, Editors. 1996, NHS Centre for Reviews and Dissemination, University of York: York.
280. Stubbs, P., *'Ethnically sensitive' or 'anti-racist'? Models for health research and service delivery*, in *'Race' and health in contemporary Britain*, W.I.U. Ahmad, Editor. 1993, Open University Press: Buckingham.
281. Lipsedge, M., *Mental health: access to care for black and ethnic minority group people.*, in *Access to health care for ethnic minorities*, A. Hopkins and V. Bahl, Editors. 1993, Royal College of General Practitioners: London.
282. Cochrane, R. and S.S. Bal, *Mental hospital admission rates of immigrants to England: a comparison of 1971 and 1981.* Soc Psychiatry Psychiatr Epidemiol, 1989. **24**(1): p. 2-11.
283. Iley, K. and J. Nazroo, *Sociology of health and illness*, in *Culture and mental health: a comprehensive textbook*, K. Bhui and D. Bhugra, Editors. 2007, Hodder Arnold: London.
284. Hillier, S. and S. Rahman, *Childhood development and behavioural and emotional problems as perceived by Bangladeshi parents in East London*, in *Researching cultural differences in health*, D. Kelleher and S. Hillier, Editors. 1996, Routledge: London.
285. Bulmer, M. and J. Solomos, *Introduction: Re-thinking Ethnic and Racial Studies* Ethnic and Racial Studies, 1998. **21**(5): p. 819-837.
286. Krieger, N., *Refiguring "race": epidemiology, racialized biology, and biological expressions of race relations.* Int J Health Serv, 2000. **30**(1): p. 211-6.
287. Sheldon, T.A. and H. Parker, *Race and ethnicity in health research.* J Public Health Med, 1992. **14**(2): p. 104-10.

288. Kertzer, D.I. and D. Arel, *Censuses, identity formation, and the struggle for political power*, in *Census and identity: the politics of race, ethnicity and language in national censuses*, D.I. Kertzer and D. Arel, Editors. 2002, Cambridge University Press: Cambridge, UK.
289. Nazroo, J.Y., *Patterns of and explanations for ethnic inequalities in health*, in *Explaining Ethnic Differences*, D. Mason, Editor. 2003, The Policy Press: Bristol.
290. Bhopal, R., *Ethnicity and race as epidemiological variables: centrality of purpose and context*, in *Health and Ethnicity*, H. Macbeth and S. Prakash, Editors. 2001, Taylor & Francis: London.
291. Fenton, S. and K. Charsley, *Epidemiology and Sociology as Incommensurate Games: Accounts from the Study of Health and Ethnicity.* Health, 2000. **4**(4): p. 403-425.
292. Lee, C., *"Race" and "ethnicity" in biomedical research: How do scientists construct and explain differences in health?* Social Science & Medicine, 2009. **68**(6): p. 1183-1190.
293. Comstock, R.D., E.M. Castillo, and S.P. Lindsay, *Four-year review of the use of race and ethnicity in epidemiologic and public health research.* Am J Epidemiol, 2004. **159**(6): p. 611-9.
294. Smart, A., et al., *The standardisation of race and ethnicity in biomedical science editorials and UK biobanks.* Soc Stud Sci, 2008. **38**(3): p. 407-423.
295. Peach, C., *South Asian migration and settlement in Great Britain, 1951-2001.* Contemporary South Asia, 2006. **15**(2): p. 133-146.
296. Department for Education and Skills. *Ethnicity and evidence. The evidence on minority ethnic pupils aged 5-16*. 2006 13th December 2006 from http://www.dfes.gov.uk/research/data/uploadfiles/0208-2006DOM-EN.pdf].
297. Kaufman, J.S. and R.S. Cooper, *Seeking causal explanations in social epidemiology.* Am J Epidemiol, 1999. **150**(2): p. 113-20.
298. Holland, P.W., *Statistics and causal inference.* Journal of American Statistical Association, 1986. **81**: p. 945-960.
299. Rutter, M., *Natural experiments, causal influences, and policy development.*, in *Ethnicity and Causal Mechanisms*, M. Rutter and M. Tienda, Editors. 2005, Cambridge University Press: Cambridge.
300. Ahmad, W.I.U., *The trouble with culture*, in *Researching cultural differences in health*, D. Kelleher and S. Hillier, Editors. 1996, Routledge: London.
301. Ellison, G.T., et al., *Racial categories in medicine: a failure of evidence-based practice?* PLoS Med, 2007. **4**(9): p. e287.
302. La Veist, T.A., *Why we should continue to study race...but do a better job: an essay on race, racism and health.* Ethn Dis, 1996. **6**(1-2): p. 21-9.
303. Davey Smith, G., *Learning to live with complexity: ethnicity, socioeconomic position, and health in Britain and the United States.* Am J Public Health, 2000. **90**(11): p. 1694-8.
304. Nazroo, J.Y., *The structuring of ethnic inequalities in health: economic position, racial discrimination, and racism.* Am J Public Health, 2003. **93**(2): p. 277-84.
305. Acheson, D., *Independent Inquiry into Inequalities in Health*. 1998, London: HMSO.
306. Galobardes, B., et al., *Indicators of socioeconomic position (part 2).* J Epidemiol Community Health, 2006. **60**(2): p. 95-101.

307. Nazroo, J.Y., *The Health of Britain's Ethnic Minorities: Findings from a National Survey*. 1997, London: Policy Studies Institute.
308. Vistram, R., *Ayahs, Lascars and Princes: Indians in Britain 1700-1947*. 1987, London: Pluto Press.
309. Peach, C., *Ethnicity in the 1991 census: Volume 2, The Ethnic Minority Population of Great Britain*. 1996, London: HMSO.
310. Ballard, R., *Desh Pardesh : the South Asian presence in Britain*. 1994, London C. Hurst.
311. Scott, A., D. Pearce, and P. Goldblatt, *The sizes and characteristics of the minority populations of Great Britain - latest estimates*, in *Population Trends, 105*, The Stationary Office, Editor. 2001: London.
312. General Register Office for Scotland. *Scotland's Census 2001: 2001 Population report*. 2002 23rd December 2006 from http://www.gro-scotland.gov.uk/files/01poprep.pdf].
313. National Statistics. *Sex and age by ethnic group (Table S101)*. 2003 8th August 2008, from http://www.statistics.gov.uk/CCI/SearchRes.asp?term=s101&x=9&y=11].
314. National Statistics, *Population by ethnic group, April 2001*. 2004, Accessed 29th January 2009 from http://www.statistics.gov.uk/statbase/Product.asp?vlnk=10991.
315. National Statistics, *Regional distribution of the minority ethnic population, April 2001*. 2001.
316. Social Exclusion Unit, *Minority ethnic issues in social exclusion and neighbourhood renewal*, UK Cabinet Office, Editor. 2000, London,.
317. Simpson, L., et al., *Ethnic minority populations and the labour market: an analysis of the 1991 and 2001 Census*. 2006, Department for Work and Pensions, Research Report No 333: London.
318. Hiebert, D. and D. Ley, *Assimilation, Cultural Pluralism, and Social Exclusion among Ethnocultural Groups in Vancouver.* Urban Geography, 2003. **24**(1): p. 16-44.
319. Peach, C., *Urban social segregation*. 1975, London: Longman.
320. Gans, H.J., *Toward a Reconciliation of "Assimilation" and "Pluralism": The Interplay of Acculturation and Ethnic Retention.* International Migration Review, 1997. **31**(4): p. 875-892.
321. Boal, F.W., *Ethnic residential segregation*, in *Social Areas in Cities, Vol. 1.*, D. Herbert and R.J. Johnston, Editors. 1976, Wiley: Chichester. p. 41-78.
322. National Statistics. *Social Trends*. 2002 7th December 2006 from http://www.statistics.gov.uk/downloads/theme_social/Social_Trends32/Social_Trends32.pdf].
323. Platt, L., *Poverty and ethnicity in the UK*. 2007, The Policy Press Oxford.
324. National Statistics, *Highest qualification: by sex and ethnic group, 2004.* . 2004, Accessed 29th January 2009 from http://www.statistics.gov.uk/statbase/Product.asp?vlnk=10991.
325. Kelly, Y., et al., *Why does birthweight vary among ethnic groups in the UK? Findings from the Millennium Cohort Study.* J Public Health (Oxf), 2008.
326. National Statistics. *Age and highest level of qualification by ethnic group (Table S117)*. 2003 13th March 2009]; Available from: http://www.statistics.gov.uk/StatBase/ssdataset.asp?vlnk=7532&Pos=&ColRank=2&Rank=256.

327. Department for Work and Pensions. *Households below average income 2000/01*. 2002 7th December 2006 from http://www.dwp.gov.uk/asd/hbai/hbai2001/pdfs/ch4tabs/tab4_1.xls].
328. National Statistics (2009) *Housing*.
329. Bhopal, R., et al., *Heterogeneity of coronary heart disease risk factors in Indian, Pakistani, Bangladeshi, and European origin populations: cross sectional study.* Bmj, 1999. **319**(7204): p. 215-20.
330. Bhattacharyya, G., L. Ison, and M. Blair, *Minority Ethnic Attainment and Participation in Education and Training: The Evidence. Research Topic Paper RTP01-03*. 2003, Department for Education and Skills: London.
331. Department for Education and Skills, *National Curriculum Assessment, GCSE and Equivalent Attainment and Post-16 Attainment by Pupil Characteristics in England 2004*. 2005, Statistical First Release. Accessed 29th January 2009 from http://www.statistics.gov.uk/cci/nugget.asp?id=1699.
332. DfES, *Ethnicity and education: The evidence on minority ethnic pupils (Research Topic Paper: RTP01-05)*. 2005, Department for Education and Skills: Nottingham.
333. Dosanjh, J.S. and P.A.S. Ghuman, *Asian Parents and English Education -- 20 years on: a study of two generations.* Educational Studies, 1997. **23**(3): p. 459 - 471.
334. Abbas, T., *The Impact of Religio-cultural Norms and Values on the Education of Young South Asian Women.* British Journal of Sociology of Education, 2003. **24**(4): p. 411-428.
335. Francis, B. and L. Archer, *British-Chinese pupils' and parents' constructions of the value of education.* British Educational Research Journal, 2005. **31**(1): p. 89 - 108.
336. Moon, N. and C. Ivins, *Parental Involvement in Children's Education*. 2004, Nottingham: Department for Education and Skills.
337. Peters, M., et al., *Parental Involvement in Children's Education 2007*. 2007: Department for Children Schools and Families.
338. National Statistics, *Age standardised 'not good health' rates: by ethnic group and sex, April 2001*. 2004, Accessed 29th January 2009 from http://www.statistics.gov.uk/statbase/Product.asp?vlnk=10991.
339. Storkey, M., *London's ethnic minorities : one city many communities*. 1994, London: London Research Centre
340. Davey Smith, G., et al., *Ethnic inequalities in health: a review of UK epidemiological evidence.* Critical Public Health, 2000. **10**(4): p. 375-408.
341. Saxena, S., J. Eliahoo, and A. Majeed, *Socioeconomic and ethnic group differences in self reported health status and use of health services by children and young people in England: cross sectional study.* Bmj, 2002. **325**(7363): p. 520.
342. Cooper, H., C. Smaje, and S. Arber, *Use of health services by children and young people according to ethnicity and social class: secondary analysis of a national survey.* Bmj, 1998. **317**(7165): p. 1047-51.
343. Nazroo, J., et al., *Children's Health*, in *Health Survey for England: the Health of Minority Ethnic Groups, '99*, Office for National Statistics, Editor. 1999, Stationery Office: London.
344. Whincup, P.H., et al., *Early evidence of ethnic differences in cardiovascular risk: cross sectional comparison of British South Asian and white children.* Bmj, 2002. **324**(7338): p. 635.

345. Dex, S. and H. Joshi, *Millennium Cohort Study: First Survey: A User's Guide to Initial Findings*. 2004, London: Centre for Longitudinal Studies, University of London.
346. Viner, R.M., et al., *Variations in associations of health risk behaviors among ethnic minority early adolescents.* J Adolesc Health, 2006. **38**(1): p. 55.
347. Smith, D.J., *Explaining ethnic variations in crime and antisocial behavior in the UK.*, in *Ethnicity and Causal Mechanisms*, M. Rutter and M. Tienda, Editors. 2005, Cambridge University Press: Cambridge.
348. Home Office, *2003 Home Office Citizenship Survey: People, Families and Communities*. 2004, London: Home Office Research, Development and Statistics Directorate.
349. Hall, K., *Lives in translation : Sikh youth as British citizens*. 2002, Philadelphia: University of Pennsylvania Press.
350. Bhugra, D. and V. Bahl, *Ethnicity : an agenda for mental health*. 1999, London: Gaskell.
351. Dolan, B., *Cross-cultural aspects of anorexia nervosa and bulimia: a review.* International Journal of Eating disorders, 1991. **10**: p. 67-78.
352. Dwivedi, K. and V. Varma, *Meeting the needs of ethnic minority children : a handbook for professionals*. 2nd ed. 2002, London Jessica Kingsley.
353. Kramer, T. and M. Hodes, *The mental health of British Afro-Caribbean children and adolescents*, in *Main issues in mental health and race*, D. Ndegwa and D. Olajide, Editors. 2003, Ashgate: Aldershot.
354. London, M., *Mental illness among immigrant minorities in the United Kingdom.* Br J Psychiatry, 1986. **149**(265-273).
355. Madge, N., *Understanding difference. The meaning of ethnicity for young lives.* 2001, London: National Children's bureau.
356. Messent, P., *Ethnic Minority Children and Families and Mental Health: Preventive approaches*, in *Promoting the emotional well-being of children and adolescents and preventing their mental ill health: A handbook.* , K. Dwivedi and P. Harper, Editors. 2004, Jessica Kingsley Publishers: London.
357. Nicol, A.R., *The problems faced by young immigrants.* Practitioner 1974. **213**(1275): p. 329-34.
358. Rutter, M. and M. Tienda, *Ethnicity and Causal Mechanisms*. 2005, Cambridge: Cambridge University Press.
359. Scott, A., M. Shaw, and C. Joughin, *The mental health of children and adolescents from ethnic minorities.*, in *Finding the evidence: a gateway to the literature in child and adolescent mental health*. 2001, Gaskell/Royal college of psychiatrists: London.
360. Smaje, C., *Health 'Race' and ethnicity: making sense of the evidence*. 1996, London: King's Fund Institute.
361. Thompson, N. and D. Bhugra, *Rates of deliberate self-harm in Asians: Findings and models.* International Review of Psychiatry, 2000. **12**(1): p. 37-43.
362. Malek, M., *Meeting the needs of minority ethnic groups in the UK*, in *Mental Health Services for Minority Ethnic Children and Adolescents*, M. Malek and C. Joughin, Editors. 2004, Jessica Kingsley: London.
363. Barn, R., *Black youth on the margins: a research review*. 2001, Joseph Rowntree Foundation,.

364. Goodman, A., V. Patel, and D.A. Leon, *Child mental health differences amongst ethnic groups in Britain: a systematic review*. BMC Public Health, 2008. **8**(1): p. 258.
365. Stroup, D.F., et al., *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.* Jama, 2000. **283**(15): p. 2008-12.
366. Verhulst, F.C., *A review of community studies*, in *The Epidemiology of Child and Adolescent Psychopathology*, F.C. Verhulst and M. Koot, Editors. 1995, Oxford University Press: New York. p. 146-177.
367. Roberts, R.E., C.C. Attkisson, and A. Rosenblatt, *Prevalence of psychopathology among children and adolescents.* Am J Psychiatry, 1998. **155**(6): p. 715-25.
368. National Statistics. *The Classification of Ethnic Groups*. 2001 8th August 2008, from http://www.statistics.gov.uk/about/Classifications/ns_ethnic_classification.asp].
369. Bhugra, D., et al., *Deliberate self-harm in adolescents in West London: Socio-cultural factors.* European Journal of Psychiatry, 2004. **18**(2): p. 91-98.
370. Ramirez, A.J., et al., *Factors predicting delayed presentation of symptomatic breast cancer: a systematic review*. Lancet, 1999. **353**(9159): p. 1127-31.
371. De Silva, M.J., et al., *Social capital and mental illness: a systematic review*. J Epidemiol Community Health, 2005. **59**(8): p. 619-27.
372. Osborn, A.F. and N.R. Butler, *Ethnic minority children : a comparative study from birth to five years : a report of the Child Health and Education Study (1970 Birth Cohort)*. 1985, London: Commission for Racial Equality.
373. Cochrane, R., *Psychological and behavioural disturbance in West Indians, Indians and Pakistanis in Britain: a comparison of rates among children and adults.* Br J Psychiatry, 1979. **134**: p. 201-10.
374. Bagley, C., *Deviant behaviour in English and West Indian schoolchildren.* Research in Education, 1972. **8**: p. 47-55.
375. Kallarackal, A.M. and M. Herbert, *The happiness of Indian immigrant children.* New Society, 1976. **26 February**: p. 422-424.
376. Earls, F. and N. Richman, *The prevalence of behavior problems in three-year-old children of West Indian-Born parents.* J Child Psychol Psychiatry 1980. **21**(2): p. 99-106.
377. Hackett, L., R. Hackett, and D.C. Taylor, *Psychological disturbance and its associations in the children of the Gujarati community*. J Child Psychol Psychiatry, 1991. **32**(5): p. 851-6.
378. Mumford, D.B., A.M. Whitehouse, and M. Platts, *Sociocultural correlates of eating disorders among Asian schoolgirls in Bradford.* Br J Psychiatry, 1991. **158**: p. 222-8.
379. Newth, S.J. and J. Corbett, *Behaviour and emotional problems in three-year-old children of Asian parentage.* J Child Psychol Psychiatry, 1993. **34**(3): p. 333-52.
380. Ahmad, S., G. Waller, and C. Verduyn, *Eating attitudes among Asian schoolgirls: the role of perceived parental control.* Int J Eat Disord, 1994. **15**(1): p. 91-7.
381. McCourt, J. and G. Waller, *Developmental role of perceived parental control in the eating psychopathology of Asian and Caucasian schoolgirls.* Int J Eat Disord, 1995. **17**(3): p. 277-82.
382. Waller, G., M. Coakley, and L. Richards, *Bulimic attitudes among Asian and Caucasian schoolgirls.* European Eating Disorders Review, 1995. **3**(1): p. 24-34.

383. Marks, F.M., *Symptoms in children of British and of Bangladeshi parents measured by the Rutter B2 Questionnaire*. Ethn Health, 1997. **2**(3): p. 255-9.
384. Furnham, A. and S. Adam-Saib, *Abnormal eating attitudes and behaviours and perceived parental control: a study of white British and British-Asian school girls.* Soc Psychiatry Psychiatr Epidemiol, 2001. **36**(9): p. 462-70.
385. Thomas, C.L., A.C. James, and M.O. Bachmann, *Eating attitudes in English secondary school students: Influences of ethnicity, gender, mood, and social class.* International Journal of Eating Disorders, 2002. **31**(1): p. 92-96.
386. Bhugra, D. and K. Bhui, *Eating disorders in teenagers in East London: A survey.* European Eating Disorders Review, 2003. **11**(1): p. 46-57.
387. Atzaba-Poria, N., A. Pike, and M. Barrett, *Internalising and externalising problems in middle childhood: A study of Indian (ethnic minority) and English (ethnic majority) children living in Britain.* International Journal of Behavioral Development, 2004. **28**(5): p. 449-460.
388. Stansfeld, S.A., et al., *Ethnicity, social deprivation and psychological distress in adolescents: school-based epidemiological study in east London.* Br J Psychiatry, 2004. **185**: p. 233-8.
389. Edmunds, S., et al., *Child Health Assessment at School Entry (CHASE) project: evaluation in 10 London primary schools.* Child Care Health Dev, 2005. **31**(2): p. 143-54.
390. Frosh, S., et al., *Prevalence of Emotional and Behavioural Disorders Among Strictly Orthodox Jewish Children in London.* Clinical Child Psychology and Psychiatry, 2005. **10**(3): p. 351-368.
391. Flouri, E., *Father's Involvement and Psychological Adjustment in Indian and White British Secondary School Age Children.* Child and Adolescent Mental Health, 2005. **10**(1): p. 32-39.
392. Berry, J., et al., *Immigrant Youth in Cultural Transition. Acculturation, Identity, and Adaptation Across National Contexts.* 2006, Mahwah, NJ, US: Lawrence Erlbaum Associates.
393. Fuller, E., *Children's health*, in *The health of minority ethnic groups, Volume 1*, K. Sproston and J. Mindell, Editors. 2006, The Information Centre: Leeds.
394. Scott, S., T. O'Connor, and A. Futh, *What makes parenting programmes work in disadvantaged areas? : the PALS trial*. 2006, York Joseph Rowntree Foundation.
395. Graham, P.J. and C.E. Meadows, *Psychiatric disorder in the children of West Indian immigrants.* J Child Psychol Psychiatry, 1967. **8**(2): p. 105-16.
396. Taylor, E.A. and S.A. Stansfeld, *Children who poison themselves. I. A clinical comparison with psychiatric controls.* Br J Psychiatry, 1984. **145**: p. 127-32.
397. Glover, G., F. Marks, and M. Nowers, *Parasuicide in young Asian women.* British Journal of Psychiatry., 1989. **154**: p. 271-272.
398. Stern, G., D. Cottrell, and J. Holmes, *Patterns of attendance of child psychiatry out-patients with special reference to Asian families.* Br J Psychiatry, 1990. **156**: p. 384-7.
399. Jawed, S.H., *A survey of psychiatrically ill Asian children.* Br J Psychiatry, 1991. **158**: p. 268-70.
400. McGibben, L., et al., *Deliberate self-poisoning in Asian and Caucasian 12-15-year-olds.* Br J Psychiatry, 1992. **161**: p. 110-2.

401. Goodman, R. and H. Richards, *Child and adolescent psychiatric presentations of second-generation Afro-Caribbeans in Britain.* Br J Psychiatry, 1995. **167**(3): p. 362-9.
402. Roberts, N. and D. Cawthorpe, *Immigrant child and adolescent psychiatric referrals: a five-year retrospective study of Asian and Caucasian families.* Can J Psychiatry, 1995. **40**(5): p. 252-6.
403. Goddard, N., F. Subotsky, and E. Fombonne, *Ethnicity and adolescent deliberate self-harm.* J Adolesc, 1996. **19**(6): p. 513-21.
404. Kramer, T., N. Evans, and M.E. Garralda, *Ethnic diversity among Child and Adolescent (CAP) clinic attenders.* Child Psychology and Psychiatry, 2000. **5**(4): p. 169-175.
405. Jayarajan, U., *The Demographic Profile of the Children and Young People Referred to and Seen by Birmingham CAMHS. Unpublished report.* 2001.
406. Lamb, G., A. Anfield, and A. Sheeran, *Access to a child mental health service. A comparison of Bangladeshi and non-Bangladeshi families.* Psychiatric Bulletin, 2002. **26**: p. 15-18.
407. Willis, J., *Audit of the Ethnicity of Referrals to a Child and Adolescent Mental Health*, in *South Asian Child and Adolescent Mental Health Service Development. Final Report*. 2004, Unpublished report from the Manchester South Asian Child and Adolescent Mental Health Project: Manchester.
408. Bhugra, D., et al., *Inception rates of deliberate self-harm among adolescents in West London.* Int J Soc Psychiatry, 2003. **49**(4): p. 247-50.
409. Messent, P. and M. Murrell, *Research leading to action: a study of accessibility of a CAMH service to ethnic minority families.* Child and Adolescent Mental Health, 2003. **8**(3): p. 118-124.
410. Minnis, H., et al., *Cultural and language mismatch: clinical complications.* Clinical child psychology and psychiatry, 2003. **8**(2): p. 179-186.
411. Tolmac, J. and M. Hodes, *Ethnic variation among adolescent psychiatric in-patients with psychotic disorders.* Br J Psychiatry, 2004. **184**: p. 428-31.
412. Hackett, L. and J. Patel, *Audit of Child and Adolescent Psychiatry referrals to the Winnicott centre*, in *South Asian Child and Adolescent Mental Health Service Development. Final Report.* 2004, Unpublished report from the Manchester South Asian Child and Adolescent Mental Health Project: Manchester.
413. Maynard, M.J. and S. Harding, *Ethnic differences in psychological well-being in adolescence in the context of time spent in family activities.* Soc Psychiatry Psychiatr Epidemiol., 2009.
414. Maynard, M.J., S. Harding, and H. Minnis, *Psychological well-being in Black Caribbean, Black African, and White adolescents in the UK Medical Research Council DASH study.* Soc Psychiatry Psychiatr Epidemiol, 2007. **42**(9): p. 759-69.
415. Rutter, M., et al., *Children of West Indian immigrants. III. Home circumstances and family patterns.* J Child Psychol Psychiatry, 1975. **16**(2): p. 105-23.
416. Ford, T., R. Goodman, and H. Meltzer, *The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders.* J Am Acad Child Adolesc Psychiatry, 2003. **42**(10): p. 1203-11.
417. Ford, T., R. Goodman, and H. Meltzer, *The relative importance of child, family, school and neighbourhood correlates of childhood psychiatric disorder.* Soc Psychiatry Psychiatr Epidemiol, 2004. **39**(6): p. 487-96.

418. Evans, R., *Ethnic differences in ADHD and the mad/bad debate.* Am J Psychiatry 2004. **161**(5): p. 932; author reply 932.
419. Klineberg, E., *Social support, ethnicity and mental health in adolescents.* Soc Psychiatry Psychiatr Epidemiol., 2006. **41**(9): p. 755-760.
420. Atzaba-Poria, N. and A. Pike, *Why do ethnic minority (Indian) children living in Britain display more internalizing problems than their English peers? The role of social support and parental style as mediators.* International Journal of Behavioral Development., 2005. **29**(6): p. 532-540.
421. Atzaba-Poria, N. and A. Pike, *Are ethnic minority adolescents at risk for problem behaviour? Acculturation and intergenerational acculturation discrepancies in early adolescence.* British Journal of Developmental Psychology, 2007. **25**(4): p. 527-541.
422. Laurens, K.R., et al., *Psychotic-like experiences and other antecedents of schizophrenia in children aged 9-12 years: a comparison of ethnic and migrant groups in the United Kingdom.* Psychol Med, 2007. **15**: p. 1-9.
423. Lewis, G., et al., *Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers.* Psychol Med, 1992. **22**(2): p. 465-86.
424. Brugha, T., et al., *Risk factors and the prevalence of neurosis and psychosis in ethnic groups in Great Britain.* Soc Psychiatry Psychiatr Epidemiol, 2004. **39**(12): p. 939-46.
425. Aaron, R., et al., *Suicides in young people in rural southern India.* Lancet, 2004. **363**(9415): p. 1117-8.
426. Pillai, A., et al., *Non-traditional lifestyles and the prevalence of mental disorders in young adolescents in a community in India.* British Journal of Psychiatry, 2008. **192**: p. 45-51.
427. Hackett, R., et al., *The prevalence and associations of psychiatric disorder in children in Kerala, South India.* J Child Psychol Psychiatry, 1999. **40**(5): p. 801-7.
428. Malhotra, S., A. Kohli, and P. Arun, *Prevalence of psychiatric disorders in school children in Chandigarh, India.* Indian J Med Res, 2002. **116**: p. 21-8.
429. Srinath, S., et al., *Epidemiological study of child & adolescent psychiatric disorders in urban & rural areas of Bangalore, India.* Indian J Med Res, 2005. **122**(1): p. 67-79.
430. Bhui, K., et al., *Cultural identity, clothing and common mental disorder: a prospective school-based study of white British and Bangladeshi adolescents.* J Epidemiol Community Health, 2008. **62**(5): p. 435-441.
431. Leon, D.A., et al., *Cohort profile: the Aberdeen children of the 1950s study.* Int J Epidemiol, 2006. **35**(3): p. 549-52.
432. Shepherd, M., B. Oppenheim, and S. Mitchell, *Childhood behaviour and mental health*. 1971, London University of London Press.
433. Rutter, M., et al., *Attainment and adjustment in two geographical areas. I--The prevalence of psychiatric disorder.* Br J Psychiatry, 1975. **126**: p. 493-509.
434. Leslie, S.A., *Psychiatric disorder in the young adolescents of an industrial town.* Br J Psychiatry, 1974. **125**(0): p. 113-24.
435. Mensah, F.K. and J. Hobcraft, *Childhood deprivation, health and development: associations with adult health in the 1958 and 1970 British prospective birth cohort studies.* J Epidemiol Community Health, 2008. **62**(7): p. 599-606.
436. Davie, R., H. Butler, and H. Goldstein, *From birth to seven: the second report of the National Child Development Study (1958 cohort)*. 1972, London: Longman.

437. Buchanan, A., J. Ten Brinke, and E. Flouri, *Parental background, social disadvantage, public "care," and psychological problems in adolescence and adulthood.* J Am Acad Child Adolesc Psychiatry, 2000. **39**(11): p. 1415-23.
438. Stallard, P., *The behaviour of 3-year-old children: prevalence and parental perception of problem behaviour: a research note.* J Child Psychol Psychiatry, 1993. **34**(3): p. 413-21.
439. McMunn , A., et al., *Psychological Health (Chapter 10)*, in *Health Survey for England: The health of young people '95-97*, P. Prescott-Clarke and P. Primatesta, Editors. 1998, The Stationery Office: London.
440. Sproston, K. and P. Primatesta, *Health Survey for England 2002: The Health of Children and Young People. Volume 3: Methodology and Documentation*. 2003, London: The Stationery Office.
441. Hobcraft, J., *Intergenerational and life-course transmission of social exclusion: influences of childhood poverty, family disruption, and contact with the police: CASEpaper 15*. 1998, LSE Centre for Analysis of Social Exclusion.
442. Fogelman, K., *Britain's sixteen year olds : preliminary findings of the third follow-up of the National Child Development Study (1958 cohort)*. 1976, London National Children's Bureau.
443. Fogelman, K., *Growing up in Great Britain : papers from the National Child Development Study*. 1983, London: Macmillan for the National Children's Bureau.
444. National Statistics. *Neighbourhood statistics topics; Child Benefit* 2006 [cited 23rd December 2006, from http://neighbourhood.statistics.gov.uk/dissemination/datasetList.do?Expand4=1&$ph=60&updateRequired=true&step=1&CurrentTreeIndex=-2#4.
445. Parry-Langdon et al, *Three years on: Survey of the development and emotional well-being of children and young people*. 2008, Cardiff: Office for National Statistics.
446. Ford, T., et al., *Child mental health is everybody's business; the prevalence of contacts with public sectors services by the types of disorder among British school children in a three-year period.* Child and Adolescent Mental Health, 2007. **12**: p. 13-20.
447. Daryanani, R., et al., *Ethnicity and use of a child and adolescent mental health service.* Child Psychology and Psychiatry Review., 2001. **6**(3): p. 127-132.
448. Goodman, R., *Psychometric properties of the strengths and difficulties questionnaire.* J Am Acad Child Adolesc Psychiatry, 2001. **40**(11): p. 1337-45.
449. Goodman, R., *A modified version of the Rutter parent questionnaire including extra items on children's strengths: a research note.* J Child Psychol Psychiatry, 1994. **35**(8): p. 1483-94.
450. Bettge, S., et al., *[Methodological comparison between the Child Behavior Checklist and the Strengths and Difficulties Questionnaires].* Gesundheitswesen, 2002. **64 Suppl 1**: p. S119-24.
451. Goodman, R. and S. Scott, *Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful?* J Abnorm Child Psychol, 1999. **27**(1): p. 17-24.
452. Goodman, R., *The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden.* J Child Psychol Psychiatry, 1999. **40**(5): p. 791-9.

453. Truman, J., et al., *The Strengths and Difficulties Questionnaire: a pilot study of a new computer version of the self-report scale.* Eur Child Adolesc Psychiatry, 2003. **12**(1): p. 9-14.
454. Goodman, R., D. Renfrew, and M. Mullick, *Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka.* Eur Child Adolesc Psychiatry, 2000. **9**(2): p. 129-34.
455. Goodman, R., et al., *Using the Strengths and Difficulties Questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders.* Eur Child Adolesc Psychiatry, 2004. **13 Suppl 2**: p. II25-31.
456. Goodman, A. and R. Goodman, *Strengths and difficulties questionnaire as a dimensional measure of child mental health.* J Am Acad Child Adolesc Psychiatry, 2009. **48**(4): p. 400-3.
457. Ford, T., et al., *A prospective study of childhood psychopathology: independent predictors of change over three years.* Soc Psychiatry Psychiatr Epidemiol, 2007. **42**(12): p. 953-61.
458. Goodman, R., et al., *Child, family, school and neighbourhood factors and their associations with childhood psychopathology: which are the independent predictors?* Unpublished.
459. Goodman, R., et al., *The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology.* J Child Psychol Psychiatry, 2000. **41**(5): p. 645-55.
460. Taylor, E., et al., *Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients.* Br J Psychiatry, 1986. **149**: p. 760-7.
461. Koskelainen, M., A. Sourander, and M. Vauras, *Self-reported strengths and difficulties in a community sample of Finnish adolescents.* Eur Child Adolesc Psychiatry, 2001. **10**(3): p. 180-5.
462. Dickey, W. and S. Blumberg, *Revisiting the Factor Structure of the Strengths and Difficulties Questionnaire: United States, 2001.* Journal of the American Academy of Child and Adolescent Psychiatry, 2004. **43**(9): p. 1159-1167.
463. Van Leeuwen, K., et al., *The Strengths and Difficulties Questionnaire in a Community Sample of Young Children in Flanders.* European Journal of Psychological Assessment, 2006. **22**(3): p. 189-197.
464. Mellor, D. and M. Stokes, *The Factor Structure of the Strengths and Difficulties Questionnaire.* European Journal of Psychological Assessment, 2007. **23**(2): p. 105-112.
465. Ronning, J.A., et al., *The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples.* Eur Child Adolesc Psychiatry, 2004. **13**(2): p. 73-82.
466. Mullick, M.S. and R. Goodman, *The prevalence of psychiatric disorders among 5-10 year olds in rural, urban and slum areas in Bangladesh: an exploratory study.* Soc Psychiatry Psychiatr Epidemiol, 2005. **40**(8): p. 663-71.
467. Mullick, M.S. and R. Goodman, *Questionnaire screening for mental health problems in Bangladeshi children: a preliminary study.* Soc Psychiatry Psychiatr Epidemiol, 2001. **36**(2): p. 94-9.
468. Samad, L., et al., *Child and adolescent psychopathology in a developing country: testing the validity of the strengths and difficulties questionnaire (Urdu version).* Int J Methods Psychiatr Res, 2005. **14**(3): p. 158-66.

469. Fleitlich-Bilyk, B. and R. Goodman, *Prevalence of child and adolescent psychiatric disorders in southeast Brazil.* J Am Acad Child Adolesc Psychiatry, 2004. **43**(6): p. 727-34.
470. Costello, E.J., et al., *The Great Smoky Mountains Study of Youth. Goals, design, methods, and the prevalence of DSM-III-R disorders.* Arch Gen Psychiatry, 1996. **53**(12): p. 1129-36.
471. Rettew, D.C., et al., *Meta-Analyses of Diagnostic Agreement Between Clinical Evaluations and Standardized Interviews*. 2006.
472. Alyahri, A. and R. Goodman, *Validation of the Arabic Strengths and Difficulties Questionnaire and the Development and Well-Being Assessment.* East Mediterr Health J, 2006. **12 Suppl 2**: p. S138-46.
473. Boys, A., et al., *Psychiatric morbidity and substance use in young people aged 13-15 years: results from the Child and Adolescent Survey of Mental Health.* Br J Psychiatry, 2003. **182**: p. 509-17.
474. Vostanis, P., et al., *Relationship between parental psychopathology, parenting strategies and child mental health--findings from the GB national study.* Soc Psychiatry Psychiatr Epidemiol, 2006. **41**(7): p. 509-14.
475. Rowe, R., B. Maughan, and R. Goodman, *Childhood psychiatric disorder and unintentional injury: findings from a national cohort study.* J Pediatr Psychol, 2004. **29**(2): p. 119-30.
476. StataCorp, *Stata 10.0: Survey Data reference manual*. 2007, College Station, Texas: Stata Press.
477. Stapleton, L.M., *Variance estimation using replication methods in structural equation modeling with complex sample data.* Structural Equation Modeling: A Multidisciplinary Journal, 2008. **15**: p. 183-210.
478. Platt, L., L. Simpson, and B. Akinwale, *Stability and change in ethnic groups in England and Wales.* Popul Trends, 2005(121): p. 35-46.
479. Goodman, A. and R. Gatward, *Who are we missing? Area deprivation and survey participation.* Eur J Epidemiol, 2008. **23**(6): p. 379-87.
480. The American Association for Public Opinion Research, *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 2000, The American Association for Public Opinion Research: Lenexa, Kansas.
481. Mateos, P., R. Webber, and P.A. Longley, *The Cultural, Ethnic and Linguistic Classification of Populations and Neighbourhoods using Personal Names, CASA Working Paper 116*. 2007, Centre for Advanced Spatial Analysis, University College London.
482. Office of the Deputy Prime Minister. *The English Indices of deprivation 2004 (revised).* 2004 26 Feb 2008]; Available from: http://www.communities.gov.uk/documents/communities/pdf/131209.pdf.
483. National Statistics. *Census 2001 - First Results*. 2001 23rd December 2006, from http://www.statistics.gov.uk/census2001/population_data.asp].
484. Scotland's Census Results OnLine. *Table S235 SEX AND AGE BY ETHNIC GROUP*. 2007 14th November from http://www.scrol.gov.uk/scrol/common/home.jsp].
485. Beerten, R., *The effect of interviewer and area characteristics on survey response rates: an exploratory analysis.* Survey Methodology Bulletin, 1999. **45**: p. 7-15.
486. Groves, M. and R. Couper, *Nonresponse in household interview surveys*. 1998, New York: Wiley Interscience.

487. UK Childhood Cancer Study Investigators, *The United Kingdom Childhood Cancer Study: objectives, materials and methods.* Br J Cancer, 2000. **82**(5): p. 1073-102.
488. Smith, A.G., et al., *Representativeness of samples from general practice lists in epidemiological studies: case-control study.* Bmj, 2004. **328**(7445): p. 932.
489. Plewis, I. and S. Ketende, *Millennium Cohort Survey first survey: Technical report on Sampling, 3rd edition*. 2004, Centre for Longitudinal Studies, Institute of Education: London.
490. Brown, T.A., *Confirmatory Factor Analysis for Applied Research*. 2006, Guilford: Guilford Press.
491. Muthen, B., *Latent Variable Structural Equation Modeling with Categorical-Data.* Journal of Econometrics, 1983. **22**(1-2): p. 43-65.
492. Muthen, B., *A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators.* . Psychometrika, 1984. **49**(1): p. 115-132.
493. Hu, L. and P.M. Bentler, *Cut-off criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives.* Structural Equation Modeling, 1999. **6**: p. 1-55.
494. Campbell, D. and D. Fiske, *Convergent and. discriminant validation by the multitrait-multimethod matrix.* Psychological Bulletin, 1959. **56**(2): p. 81-105.
495. Nunnally, J.C. and I.H. Bernstein, *Psychometric theory*. 3rd ed. 1994, New York: McGraw Hill.
496. Streiner, D. and G. Norman, *Health Measurement Scales: A Practical Guide to their Development and Use*. 3rd ed. Oxford University Press. 2003, Oxford.
497. Sorbom, D., *A General Method for Studying Differences in Factor Means and Factor Structure* British Journal of Mathematical and Statistical Psychology, 1974. **27**(299-39).
498. Victora, C.G., et al., *The role of conceptual frameworks in epidemiological analysis: a hierarchical approach.* Int J Epidemiol, 1997. **26**(1): p. 224-7.
499. Office of the Deputy Prime Minister. *Indices of deprivation 2004 - Population Denominators*. 2007 26th Feb 2008]; Available from: http://www.communities.gov.uk/archived/general-content/communities/indicesofdeprivation/216309/indicesdeprivation/.
500. Ford, T. and R. Goodman, *The Ford score; a simple measure for predicting the level of emotional and behavioural problems within mainstream schools.* Emotional and Behavioural Difficulties, 2005. **10**(219-233).
501. Goodman, A. and T. Ford, *Validation of the Ford score as a measure for predicting the level of emotional and behavioural problems in mainstream schools.* Research in Education, 2008. **80**(1): p. 1-14.
502. OPCS, *Standard Occupational Classification, Volume 3,* . 1991, London: HMSO.
503. Rose, D. and D. Pevalin, *A researcher's guide to the national statistics socio-economic classification*. 2002, London: SAGE.
504. Rose, D., D. Pevalin, and K. O'Reilly, *The NS-SEC: Origins, Development and Use*. 2005, Basingstoke: Palgrave.
505. Goldberg, D.P. and P. Williams, *A user's guide to the General Health Questionnaire*. 1998, Windsor: NFER-Nelson
506. Furukawa, T. and D.P. Goldberg, *Cultural invariance of likelihood ratios for the General Health Questionnaire.* Lancet, 1999. **353**(9152): p. 561-2.

507. Goldberg, D.P., et al., *The validity of two versions of the GHQ in the WHO study of mental illness in general health care.* Psychol Med, 1997. **27**(1): p. 191-7.
508. Naughton, M.J. and I. Wiklund, *A critical review of dimension-specific measures of health-related quality of life in cross-cultural research.* Qual Life Res, 1993. **2**(6): p. 397-432.
509. Bandyopadhyay, G., et al., *Validity of General Health Questionnaire (GHQ-36/GHQ-12) in the psychiatric O.P.D. of a general hospital--a pilot study.* Int J Soc Psychiatry, 1988. **34**(2): p. 130-4.
510. Sriram, T.G., et al., *The General Health Questionnaire (GHQ). Comparison of the English version and a translated Indian version.* Soc Psychiatry Psychiatr Epidemiol, 1989. **24**(6): p. 317-20.
511. Bhui, K., D. Bhugra, and D. Goldberg, *Cross-cultural validity of the Amritsar Depression Inventory and the General Health Questionnaire amongst English and Punjabi primary care attenders.* Soc Psychiatry Psychiatr Epidemiol, 2000. **35**(6): p. 248-54.
512. Jacob, K.S., D. Bhugra, and A.H. Mann, *The validation of the 12-item General Health Questionnaire among ethnic Indian women living in the United Kingdom.* Psychol Med, 1997. **27**(5): p. 1215-7.
513. Werneke, U., et al., *The stability of the factor structure of the General Health Questionnaire.* Psychol Med, 2000. **30**(823-829).
514. Montazeri, A., et al., *The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version.* Health Qual Life Outcomes, 2003. **1**: p. 66.
515. Toyabe, S., et al., *Factor structure of the General Health Questionnaire (GHQ-12) in subjects who had suffered from the 2004 Niigata-Chuetsu Earthquake in Japan: a community-based study.* BMC Public Health, 2007. **7**: p. 175.
516. Penninkilampi-Kerola, V., J. Miettunen, and H. Ebeling, *A comparative assessment of the factor structures and psychometric properties of the GHQ-12 and the GHQ-20 based on data from a Finnish population-based sample.* Scand J Psychol, 2006. **47**(5): p. 431-40.
517. Campbell, A., J. Walker, and G. Farrell, *Confirmatory factor analysis of the GHQ-12: can I see that again?* Aust N Z J Psychiatry, 2003. **37**(4): p. 475-83.
518. Miller, I., et al., *The McMaster Family Assessment Device: reliability and validity.* Journal of Marital and Family Therapy, 1985. **11**: p. 345-356.
519. Byles, J., et al., *Ontario Child Health Study: reliability and validity of the general functioning subscale of the McMaster Family Assessment Device.* Fam Process, 1988. **27**(1): p. 97-104.
520. Miller, I., et al., *The McMaster Approach to Families: theory, assessment, treatment and research*
Journal of Family Therapy, 2000. **22**(2): p. 168-189.
521. Tamplin, A., I.M. Goodyer, and J. Herbert, *Family functioning and parent general health in families of adolescents with major depressive disorder.* J Affect Disord, 1998. **48**(1): p. 1-13.
522. Keitner, G.I., et al., *A cross-cultural study of family functioning.* Contemporary family therapy, 1990. **122**: p. 439-454.
523. Chen, J.L., et al., *Culturally appropriate family assessment: analysis of the Family Assessment Device in a pediatric Chinese population.* J Nurs Meas, 2003. **11**(1): p. 41-60.

524. Morris, T.M., *Culturally sensitive family assessment: an evaluation of the family assessment device used with Hawaiian-American and Japanese-American families.* Fam Process, 1990. **29**(1): p. 105-16.
525. Goddard, E. and V. Higgins, *Smoking, drinking and drug use among young teenagers in 1998, Volume I (England) and Volume II (Scotland)*. 1999, London: The Stationary Office.
526. Dunn, L.M., et al., *The British picture vocabulary scale*. 2nd ed. ed. 1997, Windsor: NFER-Nelson.
527. Elliott, C., P. Smith, and K. McCulloch, *British ability scales, second edition: Administration and scoring manual*. 1996, Windsor: NFER.
528. Liddle, E.B., M.J. Batty, and R. Goodman, *The Social Aptitudes Scale: an initial validation.* Soc Psychiatry Psychiatr Epidemiol, 2008.
529. Cox, B.D., et al., *The Health and Lifestyle Survey. Preliminary report of a nationwide survey of the physical and mental health, attitudes and lifestyle of a random sample of 9,003 British adults.* 1987, London: The Health Promotion Research Trust.
530. Krause, N., J. Liang, and V. Keith, *Personality, social support, and psychological distress in later life.* Psychol Aging, 1990. **5**(3): p. 315-26.
531. Lagakos, S.W., *Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable.* Statistics in Medicine, 1988. **7**(1-2): p. 257-274.
532. De Stavola, B.L. and D.B. Cox, *On the consequences of over-stratification.* Biometrika, 2008. **95**(4): p. 992-996.
533. Van Buuren, S., H.C. Boshuizen, and D.L. Knook, *Multiple imputation of missing blood pressure covariates in survival analysis.* Statistics in Medicine, 1999. **18**: p. 681-694.
534. Kenward, M.G. and J. Carpenter, *Multiple imputation: current perspectives.* Stat Methods Med Res, 2007. **16**(3): p. 199-218.
535. Royston, P., *Multiple imputation of missing values.* Stata Journal, 2004. **4**: p. 227-241.
536. Royston, P., *Multiple imputation of missing values: update.* Stata Journal, 2005. **5**(2): p. 188–201.
537. Phillips, A.N. and G. Davey Smith, *How independent are "independent" effects? Relative risk estimation when correlated exposures are measured imprecisely*. J Clin Epidemiol, 1991. **44**(11): p. 1223-31.
538. Maugham, B., *Educational Attainments: Ethnic Differences in the United Kingdom*, in *Ethnicity and Causal Mechanisms*, M. Rutter and M. Tienda, Editors. 2005, Cambridge University Press: Cambridge.
539. Sonuga-Barke, E. and M. Mistry, *The effect of extended family living on the mental health of three generations within two Asian communities.* Br J Clin Psychol, 2000. **39 (Pt 2)**: p. 129-41.
540. Dex, S. and K. Ward, *Parental care and employment in early childhood: analysis of the Millennium Cohort Study (MCS) sweeps 1 and 2, Research Report*. 2007, Equal Opportunities Commission: Manchester.
541. Hackett, L. and R. Hackett, *Child-rearing practices and psychiatric disorder in Gujarati and British children.* British Journal of Social Work, 1994. **24**(2): p. 191-202.
542. Ghuman, P.A.S., *Asian Adolescents in the West*. 1999, Leicester: BPS Books.

543. Bhui, K., et al., *Acculturation and health status among African-Caribbean, Bangladeshi and White British adolescents: Validation and findings from the RELACHS study*. Social Psychiatry and Psychiatric Epidemiology., 2005. **40**(4): p. 259-266.
544. Deater-Deckard, K., N. Atzaba-Poria, and A. Pike, *Mother- and father-child mutuality in Anglo and Indian British families: a link with lower externalizing problems.* J Abnorm Child Psychol, 2004. **32**(6): p. 609-20.
545. Maynard, M. and S. Harding, *Does family process explain favourable psychological well-being scores in adolescents from minority ethnic groups in the MRC DASH Study?* JECH 2007. **61**(Suppl.1): p. A43.
546. NatCen, *Millennium Cohort Study Sweep 3: Questionnaire documentation*. Vol. . 2008, London: National Centre for Social Research.
547. Bhachu, P., *Parental education strategies: the case of the Punjabi Sikhs in Britain.* New Community, 1985. **12**: p. 9-21.
548. Evans, G., *Educational failure and working class White children in Britain*. 2006, Basingstoke: Palgrave Macmillan.
549. Abbas, T., *The Home and the School in the Educational Achievements of South Asians.* Race Ethnicity and Education, 2002. **5**: p. 3.
550. Griffiths, C. and J. Fitzpatrick, eds. *Geographic variations in health*. 2001, The Stationery Office: London.
551. Willms, J.D., *Literacy proficiency of youth: evidence of converging socioeconomic gradients.* International Journal of Educational Research, 2003. **39**(3): p. 247-252.
552. Nuffield Foundation, *Time trends in parenting and outcomes for young people*. 2009, London: Nuffield Foundation.
553. Kleinman, A., *Patients and Healers in the Context of Culture*. 1980, Berkeley: University of California Press.
554. Laungani, P., *Cultural influences on identity and behavior: Indian and Britain*, in *Personality and Person Perception Across Cultures*, Y.-T. Lee, C.R. McCanley, and J.G. Draguns, Editors. 1999, Lawrence Erlbaum Associated Publishers: Mahwah, New Jersey.
555. Dosanjh, J.S. and P. Ghuman, *Child-rearing in ethnic minorities*. 1996, Avon: Multilingual Matters Ltd.
556. Hackett, L. and R. Hackett, *Parental ideas of normal and deviant child behaviour. A comparison of two ethnic groups.* Br J Psychiatry, 1993. **162**: p. 353-7.
557. Jayaweera, H., et al., *Millennium Cohort Study First Survey Demographic and socio-economic characteristics of ethnic minority mothers in England*. 2007, Oxford: Centre for Longitudinal Studies.
558. Centre for Longitudinal Studies. *Millennium Cohort Study*. 2009 22nd February 2009]; Available from: http://www.cls.ioe.ac.uk/studies.asp?section=000100020001.
559. Sijtsma, K., *On the use, the misuse, and the very limited usefulness of Cronbach's alpha.* Psychometrika, 2009. **74**(1): p. 107-120.
560. Lamping, D.L., et al., *Evaluation of outcomes in chronic venous disorders of the leg: development of a scientifically rigorous, patient-reported measure of symptoms and quality of life.* J Vasc Surg, 2003. **37**(2): p. 410-9.
561. Lamping, D.L., et al., *The Community-Acquired Pneumonia Symptom questionnaire: a new, patient-based outcome measure to evaluate symptoms in patients with community-acquired pneumonia.* Chest, 2002. **122**(3): p. 920-9.

562. Hilari, K., et al., *Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): evaluation of acceptability, reliability, and validity*. Stroke, 2003. **34**(8): p. 1944-50.
563. Smith, S.C., et al., *Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology*. Health Technol Assess, 2005. **9**(10): p. 1-93.
564. Shrout, P.E. and J.L. Fleiss *Intraclass Correlations: Uses in Assessing Rater Reliability*. Psychological Bulletin, 1979. **86**(2): p. 420-428.
565. Krishnamoorthy, K. and Y. Xia, *Inferences on correlation coefficients: One-sample, independent and correlated cases.* Journal of Statistical Planning and Inference, 2007. **137**(7): p. 2362-2379.
566. Dunn, G., *Design and analysis of reliability studies*. 1989, New York: Oxford University Press.
567. Myers, L. and M.J. Sirois, *Differences between Spearman Correlation Coefficients*, in *Encyclopedia of Statistical Sciences, Vol 12*, S. Kotz, et al., Editors. 2006, Wiley-Interscience: Hoboken, N.J/. p. 7901-7902.
568. Schmitt, N., *Uses and abuses of coefficient alpha.* Psychological Assessment, 1996. **8**: p. 350-353.
569. Bentler, P.A. and J.A. Woodward, *Inequalities among lower bounds to reliability: With applications to test construction and factor analysis.* Psychometrika, 1980. **45**: p. 249–267.
570. Fawcett, T., *An introduction to ROC analysis.* Pattern Recognition Letters 2006. **27**: p. 861-874.
571. Albright, J.J. and H.M. Park, *Confirmatory Factor Analysis Using Amos, LISREL, Mplus, and SAS/STAT CALIS: Technical Working Paper*. 2008, The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.
572. Asparouhov, A. and B. Muthen. *Exploratory Structural Equation Modeling*. Technical report. 2008 30 May 2009]; Available from: http://www.statmodel.com/download/EFACFA810.pdf.
573. Snook, S.C. and R.L. Gorsuch, *Principal component analysis versus common factor analysis: A Monte Carlo study.* Psychological Bulletin, 1989. **106**: p. 148-154.
574. Velicer, W.F. and D.N. Jackson, *Component Analysis versus Common Factor Analysis: Some issues in Selecting an Appropriate Procedure.* Multivariate Behavioral Research, 1990. **25**(1): p. 1-28.
575. Kaiser, H.F., *The application of electronic computers to factor analysis.* Educational and Psychological Measurement, 1960. **20**: p. 141-151.
576. Armitage, P., G. Berry, and J.N.S. Matthews, *Statistical methods in medical research*. 4th ed. 2002, Oxford: Blackwell Science.
577. Beauducel, A. and P.Y. Herzberg, *On the Performance of Maximum Likelihood versus Means and Variance Adjusted Weighted Least Squares Estimation in CFA* Structural Equation Modeling: A Multidisciplinary Journal, 2006. **13**(2): p. 186-203.
578. Millsap, R.E. and H.T. Everson, *Methodology review: Statistical approaches for assessing measurement bias.* Applied Psychological Measurement, 1993. **17**(4): p. 297-334.
579. Mellenbergh, G.J., *Item bias and item response theory.* International Journal of Educational Research, 1989. **13**: p. 127-143.

580. Drasgow, F., *Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues.* Psychological Bulletin, 1984. **95**: p. 134-135.
581. MacCallum, R.C., et al., *Sample size in Factor Analysis.* Psychological Methods, 1999. **4**(1): p. 84-99.
582. MacCallum, R.C., et al., *Sample Size in Factor Analysis: The Role of Model Error* Multivariate Behavioral Research, 2001. **36**(4): p. 611-637.
583. McCullagh, P. and J.A. Nelder, *Generalized Linear Models, Second Edition*. 1989, London Chapman and Hall.
584. Clogg, C.C., E. Petkova, and A. Haritou, *Statistical methods for comparing regression coefficients between models.* American Journal of Sociology, 1995. **100**: p. 1261-1293.
585. Raymond Paternoster, R.B.P.M.A.P., *Using the correct statistical test for the equality of regression coefficients.* Criminology, 1998. **36**(4): p. 859-866.
586. Collett, D., *Modelling binary data*. 2002, Boca Raton, FL: Chapman & Hall/CRC Press.
587. Agresti, A., *Introduction to categorical data analysis*. 2007, Hoboken, N.J.: Wiley; Chichester.
588. Ananth, C.V. and D.G. Kleinbaum, *Regression models for ordinal responses: a review of methods and applications.* Int J Epidemiol, 1997. **26**(6): p. 1323-33.
589. McCullagh, P., *Regression models for ordinal data.* Journal of the Royal Statistical Society, Series B, 1980. **42**: p. 109-142.
590. Whittingham, M.J., et al., *Why do we still use stepwise modelling in ecology and behaviour?* Journal of Animal Ecology, 2006. **75**: p. 1182-1189.
591. Miller, A.J., *Selection of subsets of regression variables.* Journal of the Royal Statistical Society A, 1984. **147**: p. 389-425.
592. Mundry, R. and C.L. Nunn, *Stepwise model fitting and statistical inference: turning noise into signal pollution.* Am Nat, 2009. **173**(1): p. 119-23.
593. Wilkinson, L., *Tests of significance in stepwise regression.* Psychological Bulletin, 1979. **86**: p. 168-174.
594. Scheaffer, R., W. Mendelhall, and L. Ott, *Elementary survey sampling*. 5th ed. 1996, Boston: Duxbury Press.
595. Cochran, W.G., *Sampling techniques*. 3rd ed. 1977, New York: Wiley.
596. Kirkwood, B.R. and J.A.C. Sterne, *Essential Medical Statistics*. 2nd ed. 2003, Oxford: Blackwell Publishing.
597. Asparouhov, T. and B. Muthén, *Multivariate statistical modeling with survey data*, in *Proceedings of the Federal Committee on Statistical Methodology (FCSM) Research Conference (available from http://www.statmodel.com/papers.shtml)*. 2005.
598. Skinner, C.J., D. Holt, and T.M.F. Smith, *Analysis of Complex Surveys*. 1989, New York: Wiley.
599. Kish, L. and M.R. Frankel, *Inference from complex samples.* Journal of the Royal Statistical Society, 1974. **36 (Series B)**: p. 1-37.
600. Rubin, D.B., *Inference and missing data (with discussion).* Biometrika, 1976. **63**: p. 581-592.
601. Molenberghs, G. and M.G. Kenward, *Missing Data in Clinical Studies*. 2007, Chicester: Wiley.

602. Vach, W. and M. Blettner, *Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables.* Am J Epidemiol, 1991. **134**(8): p. 895-907.
603. Schafer, J.L., *Analysis of incomplete multivariate data*. 1997, London: Chapman and Hall.
604. Carlin, J.B., *A new framework for managing and analyzing multiply imputed data in Stata.* Stata Journal, 2008. **8**(1): p. 49-67.
605. CEEHD. *UK Centre for Evidence in Ethnicity, Health and Diversity,*. 2007 14th July 2007]; Available from: http://www2.warwick.ac.uk/fac/med/research/csri/ethnicityhealth/.
606. CRER. *Centre for Research in Ethnic Relations*. 2007 13th July 2007]; Available from: http://www2.warwick.ac.uk/fac/soc/crer.
607. NHS. *NHS, Ethnicity and Health Specialist Library*. 2007 14th July 2007]; Available from: www.library.nhs.uk/ethnicity.
608. Confederation of Indian Organisations. *Confederation of Indian Organisations*. 2007 14th July 2007]; Available from: http://www.cio.org.uk/.
609. National Electronic Library for Health. *The mental health of children and adolescents from ethnic minorities*. 2007 14th July 2007]; Available from: http://www.nelmh.org/page_view.asp?c=15&fc=019005&did=1624.
610. King's Fund. *King's Fund*. 2007 14th July 2007]; Available from: http://www.kingsfund.org.uk/.
611. MIND. *MIND*. 2007 1st May 2007]; Available from: http://www.mind.org.uk/Information/Factsheets.
612. NEMDA. *NEMDA-The National Ethnic Minority Data Archive project*. 2007 14th July 2007]; Available from: http://www.warwick.ac.uk/~errac/nempubs.htm.
613. Bhui, K., et al., *Cultural identity, acculturation, and mental health among adolescents in east London's multiethnic community.* J Epidemiol Community Health, 2005. **59**(4): p. 296-302.
614. Burt, K.B., et al., *The prediction of disruptive behaviour disorders in an urban community sample: the contribution of person-centred analyses.* J Child Psychol Psychiatry, 2004. **45**(6): p. 1159-70.
615. Chowdhury, N.A., et al., *Ethnicity and its relevance in a seven-year admission cohort to an English national adolescent medium secure health service unit.* Crim Behav Ment Health, 2005. **15**(4): p. 261-72.
616. Earls, F. and N. Richman, *Behavior problems in pre-school children of West Indian-born parents: a re-examination of family and social factors.* J Child Psychol Psychiatry 1980. **21**(2): p. 107-17.
617. Earls, F., *Cultural and national differences in the epidemiology of behavior problems of preschool children.* Cult Med Psychiatry, 1982. **6**(1): p. 45-56.
618. Flouri, E., *Non-resident fathers' relationships with their secondary school age children: determinants and children's mental health outcomes.* J Adolesc, 2006. **29**(4): p. 525-38.
619. Gowers, S., et al., *Social and family factors in adolescent psychiatry.* J Adolesc, 1993. **16**(4): p. 353-66.
620. Handy, S., et al., *Ethnic differences in adolescent self-poisoning: a comparison of Asian and Caucasian groups.* J Adolesc, 1991. **14**(2): p. 157-62.
621. Hill, A.J. and R. Bhatti, *Body shape perception and dieting in preadolescent British Asian girls: links with eating disorders.* Int J Eat Disord, 1995. **17**(2): p. 175-83.

622. Holden, N.L. and P.H. Robinson, *Anorexia nervosa and bulimia nervosa in British blacks.* Br J Psychiatry, 1988. **152**: p. 544-9.
623. Morita, H., *Behavioural deviance in Japanese children attending school in England.* Acta Paediatr Jpn 1987. **29**(2): p. 233-40.
624. Shah, Q. and E. Sonuga-Barke, *Family structure and the mental health of Pakistani Muslim mothers and their children living in Britain.* Br J Clin Psychol, 1995. **34 (Pt 1)**: p. 79-81.
625. Stevenson, J., N. Richman, and P. Graham, *Behaviour problems and language abilities at three years and behavioural deviance at eight years.* J Child Psychol Psychiatry, 1985. **26**(2): p. 215-30.
626. Tareen, A., M. Hodes, and L. Rangel, *Non-fat-phobic anorexia nervosa in British South Asian adolescents.* Int J Eat Disord, 2005. **37**(2): p. 161-5.
627. Vostanis, P., et al., *Service utilisation by children with conduct disorders--findings from the GB National Study.* Eur Child Adolesc Psychiatry, 2003. **12**(5): p. 231-8.
628. Wardle, J., et al., *Depression in adolescent obesity: cultural moderators of the association between obesity and depressive symptoms.* Int J Obes (Lond), 2006. **30**(4): p. 634-43.
629. Boeing, L., et al., *Adolescent-onset psychosis: prevalence, needs and service provision.* Br J Psychiatry, 2007. **190**: p. 18-26.
630. Leavey, G., et al., *Psychological disorder amongst refugee and migrant schoolchildren in London.* Soc Psychiatry Psychiatr Epidemiol, 2004. **39**(3): p. 191-5.
631. Best, D., et al., *Excessive drinking and other problem behaviours among 14-16 year old schoolchildren.* Addict Behav, 2006. **31**(8): p. 1424-35.
632. Atzaba-Poria, N., A. Pike, and K. Deater-Deckard, *Do risk factors for problem behaviour act in a cumulative manner? An examination of ethnic minority and majority children through an ecological perspective.* J Child Psychol Psychiatry, 2004. **45**(4): p. 707-18.
633. Bendall, P., M. Hamilton, and N. Holden, *Eating disorders in Asian girls.* British Journal of Psychiatry, 1991. **159**: p. 441.
634. Lindsey, C., et al., *Prevalence of Emotional and Behavioural Disorders Among Strictly Orthodox Jewish Pre-School Children in London.* Clinical Child Psychology and Psychiatry, 2003. **8**(4): p. 459-472.
635. Mears, A., et al., *Characteristics of the Detained and Informal Child and Adolescent Psychiatric In-Patient Populations.* Child and Adolescent Mental Health, 2003. **8**(3): p. 131-134.
636. Ullah, P., M. Banks, and P. Warr, *Social support, social pressures and psychological distress during unemployment.* Psychological Medicine, 1985. **15**(2): p. 283-295.
637. Marchant, P., A. Hussain, and K. Hall, *Autistic Spectrum Disorders and Asian children.* British Journal of Educational Studies, 2006. **54**(2): p. 230-244.
638. Zaineb, D.A., *Differences in factors associated with deliberate self-harm in children and adolescents: A transcultural comparison*, [Study description in NRR database].
639. Vyas, I., et al., *Two cultures in one city. Similarities and differences in child rearing between native and Indian families of three-year-olds.*, in 1998, Poster presented at the 14th International Congress of the International Association for Child and Adolescent Psychiatry and Allied Professions

640. Zietlin, H., *Risk and Protective Factors in adolescents who attempt suicide: Differences across ethnic groups*, [Study description in NRR database].
641. Hodes, M., *West London adolescent psychosis study*, [Study description in NRR database].
642. Skinner, A., *An Epidemiological study of deliberate self-harm in an inner-city, multi-ethnic adolescent population*, [Study description in NRR database].
643. Muthalagu, S., *Eatinng Attitudes and Behaviours among Adolescents in East London - Relationship with Etnicity*, [Study description in NRR database].
644. Arnone, D., *Ethnic influence in the diagnosis of autism*, [Study description in NRR database].
645. Ramjee, M.J., *An investigation into the increase from 1983 to 2003 (20 years) of reported cases of psychosis in young girls of South Asian ethnicity in Tier 4 of the CAMHS on Ward 3 at Birmingham Children's Hospital* [Study description in NRR database].
646. Ahmad, S., G. Waller, and C. Verduyn, *Eating Attitudes and Body Satisfaction among Asian and Caucasian Adolescents.* Journal of Adolescence, 1994. **17**(5): p. 461-470.
647. Bhugra, D., et al., *Deliberate self-harm in adolescents. A case note study among two ethnic groups.* European Journal of Psychiatry, 2002. **16**(3): p. 145-151.
648. Ogden, J. and C. Elder, *The role of family status and ethnic group on body image and eating behavior.* International Journal of Eating Disorders, 1998. **23**(3): p. 309-315.
649. Kingsbury, S., *The Psychological and Social Characteristics of Asian Adolescent Overdose.* Journal of Adolescence, 1994. **17**(2): p. 131-135.
650. King's Fund, *Ethnic health: an introduction to ethnic health issues reading list*. 2006, London: King's Fund.
651. Fallon, T., Jr. and M. Schwab-Stone, *Determinants of reliability in psychiatric surveys of children aged 6-12.* J Child Psychol Psychiatry, 1994. **35**(8): p. 1391-408.
652. Schwab-Stone, M., et al., *Reliability of diagnostic reporting for children aged 6-11 years: a test-retest study of the Diagnostic Interview Schedule for Children-Revised.* Am J Psychiatry, 1994. **151**(7): p. 1048-54.
653. Schwab-Stone, M.E., et al., *Criterion validity of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3).* J Am Acad Child Adolesc Psychiatry, 1996. **35**(7): p. 878-88.
654. Minnis, H., E. Kelly, and H. Bradby, *The use of child psychiatry by South Asian Families in South Glawgow (cited in Malek, M., Meeting the needs of minority ethnic groups in the UK, in Mental Health Services for Minority Ethnic Children and Adolescents, M. Malek and C. Joughin, Editors. 2004, Jessica Kingsley: London).* "In press".
655. Minnis, H., *re: request for information [2]*, A. Goodman, Editor. 22nd March 2007.
656. Walker, M., *West Indian Children in school. Unpublished report*. 1968, [Cited in Nicol, A.R., Psychiatric disorder in the children of Caribbean immigrants. J Child Psychol Psychiatry, 1971. 12(4): p. 273-87].
657. Kelly, Y., *re: MCS Mental Health*, A. Goodman, Editor. 22nd May 2007.
658. Laurens, K.R., *RE: Antecedents of schizophrenia in children from different ethnic groups*, A. Goodman, Editor. 19th May 2007.

659. Truman, J., *RE: Child mental health and ethnicity - request for help locating studies*, A. Goodman, Editor. 22nd May 2007.
660. Kelsall, M., M. Dolan, and S. Bailey, *Violent incidents in an adolescent forensic unit.* Med Sci Law, 1995. **35**(2): p. 150-8.
661. O'Herlihy, A., et al., *National In-patient Child and Adolescent Psychiatry Study (NICAPS): Final Report to the Department of Health*. 2001, London: Royal College of Psychiatrists' Research Unit.
662. Schmidt, U., M. Hodes, and J. Treasure, *Early onset bulimia nervosa: who is at risk? A retrospective case-control study.* Psychol Med, 1992. **22**(3): p. 623-8.
663. Nicol, A.R., *Psychiatric disorder in the children of Caribbean immigrants.* J Child Psychol Psychiatry, 1971. **12**(4): p. 273-87.
664. Furnham, A. and R. Patel, *The eating attitudes and behaviours of Asian and British schoolgirls: a pilot study.* Int J Soc Psychiatry, 1994. **40**(3): p. 214-26.
665. Furnham, A., *RE: request for information*, A. Goodman, Editor. 16th February 2007.
666. Hillier, S., et al., *An evaluation of child psychiatric services for Bangladeshi parents.* Journal of Mental Health, 1994. **3**: p. 327-337.
667. Moran, P., et al., *Callous and unemotional traits in children and adolescents living in Great Britain.* British Journal of Psychiatry, 2008. **192**: p. 65-66.
668. Harris, R. and E. Hardman, *Developing and Evaluating Community Mental Health Services: The Bangladeshi Community, Assessment of Need.* Journal of Clinical Excellence, 2001. **3**(2): p. 69-73.
669. Bhatnagar, J., *Immigrants at school*. 1970, London: Cornmarket press.
670. Dosanjh, J.S., *A comparative study of Punjabi and English child rearing practices with special reference to lower juniors (7-9 years)* in *Psychology*. 1976, University of Nottingham: Nottingham.
671. Schools Council, *Teaching English to West Indian children : the research stage of the project*. 1970, London: Evans/Methuen Educational.
672. Stott, D.H., *The social-adjustment of children: Manual to the Bristol Social-Adjustment Guides*. 1963, London: University of London press.
673. Hawton, K., et al., *Deliberate self harm in adolescents: self report survey in schools in England.* Bmj, 2002. **325**(7374): p. 1207-11.
674. Fatimilehin, I.A. and P.G. Coleman, *Appropriate services for African-Caribbean families: views from one community.* Clinical Psychology Forum, 1998. **111**: p. 6-11.
675. Subotsky, F. and G. Berelowitz, *Consumer views at a community child guidance clinic.* Newsletter of the Association for Child Psychology and Psychiatry, 1990. **12**(3): p. 8-12.
676. Shams, M. and R. Williams, *Differences in perceived parental care and protection and related psychological distress between British Asian and non-Asian adolescents.* Journal of Adolescence, 1995. **18**: p. 329-348.
677. Mumford, D.B. and A.M. Whitehouse, *Increased prevalence of bulimia nervosa among Asian schoolgirls.* Bmj, 1988. **297**(6650): p. 718.
678. Jarman, B., *Identification of underprivileged areas.* Br Med J (Clin Res Ed), 1983. **286**(6379): p. 1705-9.
679. Carstairs, V. and R. Morris, *Deprivation: explaining differences in mortality between Scotland and England and Wales.* Bmj, 1989. **299**(6704): p. 886-9.

680. Adams, J. and M. White, *Removing the health domain from the Index of Multiple Deprivation 2004-effect on measured inequalities in census measure of health.* J Public Health (Oxf), 2006. **28**(4): p. 379-83.
681. Patel, V., *Cultural factors and international epidemiology.* Br Med Bull, 2001. **57**: p. 33-45.
682. Williams, R. and K. Hunt, *Psychological distress among British South Asians: the contribution of stressful situations and subcultural differences in the West of Scotland Twenty-07 Study*. Psychol Med, 1997. **27**(5): p. 1173-81.
683. Epstein, N. and B. Bishop, *Problem-centred systems therapy of the family*, in *Handbook of family therapy*, A. Gurman and D. Kniskern, Editors. 1981, Bruner/Mazel: New York.
684. Wenniger, W.F., W. Hagemand, and W.A. Arrindell, *Cross-national validity of dimensions of family functioning: first experiences with the Dutch version of the McMaster Family Assessment Device (FAD).* Pers Indiv Diff, 1993. **14**(6): p. 769-81.
685. Goodman, R., *Standardising BPVS-II scores (personal communication by email)*. 2007.